

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds

(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649a-5
Sequence: 49
1 LSVRLKXPVKE 11

Scoring table:

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-O=/cgn2_1/USPTO.spool/US09823649.runtat.21012003.093151.24626/app.query.fasta.1.1393
-MODEL=frimr.p2n.model -DEV=xlp
-LOOPEXT=0 -UNITS=bits -STAIR=1 -END=1 -MATRIX=blomsum62
-DB=PublishedApplications.NA.GPMT-fastap -SUFFIX=rmpb -MIMATCH=0.1
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=ptt -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649.@CGN.1.1.77.@runtat.21012003.093151.24626
-NCPU=6 -LOPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PublishedApplications.NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	95.9	2682	9	US-09-891-332A-1
2	37	75.5	696	10	US-09-974-300-7373
3	36	73.5	93	10	US-09-922-261-33
4	36	73.5	464	10	US-09-967-768A-154

5	36	73.5	1096	10	US-09-922-261-26	Sequence 26, Appl
6	36	73.5	43804	10	US-09-970-711-1	Sequence 1, Appl
7	35	71.4	1074	9	US-10-004-717-18	Sequence 18, Appl
8	34	69.4	1945	10	US-09-731-872-37	Sequence 37, Appl
9	34	69.4	1953	10	US-09-731-872-41	Sequence 41, Appl
10	34	69.4	1969	10	US-09-731-872-59	Sequence 59, Appl
11	34	69.4	2001	10	US-09-815-242-9539	Sequence 939, Ap
12	34	69.4	2031	10	US-09-815-242-9415	Sequence 9415, Ap
13	34	69.4	4668	10	US-09-286-240-5	Sequence 5, Appl
14	34	69.4	4668	10	US-09-863-777-1	Sequence 1, Appl
15	34	69.4	4668	10	US-09-880-107-2239	Sequence 2239, Ap
16	34	69.4	9388	9	US-09-924-400-141	Sequence 141, App
17	34	69.4	9388	10	US-09-810-936-141	Sequence 141, App
18	34	69.4	9388	10	US-09-429-755-141	Sequence 141, App
19	33	67.3	229	10	US-09-960-352-13394	Sequence 13394, A
20	33	67.3	257	10	US-09-878-574-5560	Sequence 5560, Ap
21	33	67.3	257	10	US-09-878-574-6474	Sequence 6474, Ap
22	33	67.3	273	10	US-09-878-574-8764	Sequence 8764, Ap
23	33	67.3	277	10	US-09-867-701-6945	Sequence 6945, Ap
24	33	67.3	390	10	US-09-960-352-4518	Sequence 4518, Ap
25	33	67.3	409	10	US-09-960-352-14105	Sequence 14106, A
26	33	67.3	868	10	US-09-770-445-567	Sequence 567, App
27	33	67.3	891	9	US-09-938-842A-1994	Sequence 1994, App
28	33	67.3	2580	9	US-09-738-626-2964	Sequence 2964, Ap
29	33	67.3	2768	10	US-09-905-983-4	Sequence 4, Appl
30	33	67.3	2768	10	US-09-905-983-6	Sequence 6, Appl
31	33	67.3	8546	10	US-09-070-927A-146	Sequence 146, App
32	33	67.3	19206	10	US-09-764-869-1258	Sequence 1258, App
33	33	67.3	19274	10	US-09-764-869-1254	Sequence 1254, Ap
34	33	67.3	31348	10	US-09-764-869-1259	Sequence 1259, Ap
35	33	67.3	659158	9	US-09-771-208-20	Sequence 20, Appl
36	32	65.3	122	10	US-09-923-876-2960	Sequence 2960, Ap
37	32	65.3	297	10	US-09-960-352-4917	Sequence 4917, Ap
38	32	65.3	363	10	US-09-294-0938-322	Sequence 322, App
39	32	65.3	363	10	US-09-770-791-477	Sequence 477, App
40	32	65.3	366	9	US-10-079-623-7	Sequence 7, Appl
41	32	65.3	428	10	US-09-960-352-4817	Sequence 4817, Ap
42	32	65.3	490	10	US-09-917-800A-1063	Sequence 1063, Ap
43	32	65.3	516	10	US-09-783-590-652	Sequence 652, App
44	32	65.3	716	9	US-09-938-842A-3318	Sequence 4318, Ap
45	32	65.3	798	9	US-09-813-453A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-891-332A-1

Sequence 1, Application US/09891332A
Patent No. US20020166646A1

GENERAL INFORMATION:

APPLICANT: Chatterjee, Deb K.
Sotus, Joseph

Yang, Shuwei

TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS: Nucleic Acid Fragments and Uses Thereof

ADDRESSER: STEINER, KESSLER, GOLDSTEIN & FOX, P. L. L. C.

STREET: 1100 New York Ave., N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/891.332A

FILING DATE: 27-Jun-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 09/019,160
; FILING DATE: <UNKNOWN>
; APPLICATION NUMBER: US 60/037,393
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Remond, Robert W
; REGISTRATION NUMBER: 32,893
; TELECOMMUNICATION INFORMATION:
; REFERENCE/DOCKET NUMBER: 0942.4250002
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-891-332A-1

Alignment Scores:
Pred. No.: 0.365 Length: 2682
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: Gaps: 0

US-09-823-649A-5 (1-11) x US-09-891-332A-1 (1-2682)
OY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 2221 CTTTCTGTGACACTTGCAATACCGTTAAAGAA 2253

RESULT 2
US-09-974-300-7373
; Sequence 7373, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085,500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7373
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-7373

Alignment Scores:
Pred. No.: 13.6 Length: 696
Score: 37.00 Matches: 7
Percent Similarity: 90.00% Conservative: 2
Best Local Similarity: 70.00% Mismatches: 1
Query Match: 75.51% Indels: 0
DB: Gaps: 0

US-09-823-649A-5 (1-11) x US-09-974-300-7373 (1-696)
OY 2 SerValArgLeuGly**ProValLysGlu 11
Db 657 TCCATGCGCCTCGTTGCGCTTAAAGAA 686
```

```

RESULT 3
US-09-922-261-33
; Sequence 33, Application US/09922261
; Patent No. US2002011471A1
; GENERAL INFORMATION:
; APPLICANT: CIGNET NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Harney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-261-33

Alignment Scores:
Pred. No.: 2.35 Length: 93
Score: 36.00 Matches: 6
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 60.00% Mismatches: 3
Query Match: 73.47% Indels: 0
DB: Gaps: 0

US-09-823-649A-5 (1-11) x US-09-922-261-33 (1-93)
OY 1 LeuSerValArgLeuGly**ProValLys 10
Db 1 ATGACGATACGACTTCGACGCCGACGTCGAG 30

RESULT 4
US-09-967-768A-154/C
; Sequence 154, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Merida
; TITLE OF INVENTION: Cancer Gene Identification and Therapeutic Screening Using Sig
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(464)
; OTHER INFORMATION: n=a,t,g or c
US-09-967-768A-154

Alignment Scores:
Pred. No.: 14.4 Length: 464
```

Score: 36.00 Matches: 7
 Percent Similarity: 81.82% Conservative: 2
 Best Local Similarity: 63.64% Mismatches: 2
 Query Match: 73.47% Indels: 0
 DB: 10 Gaps: 0

US-09-823-649A-5 (1-11) x US-09-967-768A-154 (1-464)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11
 DB 311 CTTCCGTTCGATGCGTTCACAGAAAAAG 279

RESULT 5

US-09-922-261-5
 ; Sequence 26, Application US/09922261
 ; Patient No. US20020111471A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COGENT NEUROSCIENCE, Inc.
 ; APPLICANT: Lo, Donald C.
 ; APPLICANT: Barney, Shawn
 ; APPLICANT: Thomas, Mary Beth
 ; APPLICANT: Portbury, Stuart D.
 ; APPLICANT: Putnam, Kasturi
 ; APPLICANT: Katz, Lawrence C.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
 ; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
 ; TITLE OF INVENTION: CELL DEATH
 ; FILE REFERENCE: 10001-005-999
 ; CURRENT APPLICATION NUMBER: US/09/922,261
 ; CURRENT FILING DATE: 2001-08-03
 ; PRIOR APPLICATION NUMBER: US/09/461,697
 ; PRIOR FILING DATE: 1999-12-14
 ; NUMBER OF SEQ ID NOS: 456
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 26
 ; LENGTH: 1096
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-922-261-26

Alignment Scores:
 Pred. No.: 37.9 Length: 1096
 Score: 36.00 Matches: 6
 Percent Similarity: 90.00% Conservative: 3
 Best Local Similarity: 60.00% Mismatches: 1
 Query Match: 73.47% Indels: 0
 DB: 10 Gaps: 0

US-09-823-649A-5 (1-11) x US-09-922-261-26 (1-1096)

QY 1 LeuSerValArgLeuGly***ProValLys 10
 DB 249 ATGACGATACGACGTGGGAGCCGACGTGAG 278

RESULT 6

US-09-970-711-1
 ; Sequence 1, Application US/09970711
 ; Patent No. US20020081279A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Adam
 ; APPLICANT: Colten, Matthew
 ; APPLICANT: Chioeca, Susanna
 ; APPLICANT: Kutzbauer, Robert
 ; APPLICANT: Schaffner, Gottfried
 ; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
 ; FILE REFERENCE: 0652,1800001
 ; CURRENT APPLICATION NUMBER: US/09/970,711
 ; CURRENT FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 09/171,461
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: PCT/EP97/01944
 ; PRIOR FILING DATE: 1997-04-18
 ; NUMBER OF SEQ ID NOS: 54

SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 43804
 ; TYPE: DNA
 ; ORGANISM: CELO Virus
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (12193)..(15043)
 ; OTHER INFORMATION: /gene: L1
 ; NAME/KEY: misc_feature
 ; LOCATION: (15080)
 ; OTHER INFORMATION: /note= L2 region penton base splice acceptor site
 ; NAME/KEY: gene
 ; LOCATION: (15110)..(17495)
 ; OTHER INFORMATION: /gene: L2
 ; NAME/KEY: polyA_site
 ; LOCATION: (17526)
 ; NAME/KEY: gene
 ; LOCATION: (17559)..(21754)
 ; OTHER INFORMATION: /gene: L3
 ; NAME/KEY: misc_feature
 ; LOCATION: (18261)
 ; OTHER INFORMATION: /gene: L3 /note= hexon splice acceptor site
 ; NAME/KEY: misc_feature
 ; LOCATION: (21102)
 ; OTHER INFORMATION: /gene: L3 /note= protease splice acceptor site
 ; NAME/KEY: misc_feature
 ; LOCATION: (21123)
 ; OTHER INFORMATION: /gene: L3 /note= protease splice acceptor site
 ; NAME/KEY: polyA_site
 ; LOCATION: (21767)
 ; NAME/KEY: polyA_site
 ; LOCATION: (21824)
 ; NAME/KEY: polyA_site
 ; LOCATION: (21836)
 ; NAME/KEY: polyA_site
 ; LOCATION: (21882)
 ; NAME/KEY: misc_feature
 ; LOCATION: (23608)
 ; OTHER INFORMATION: /note= 100K splice acceptor site
 ; NAME/KEY: misc_feature
 ; LOCATION: (23649)
 ; OTHER INFORMATION: /note= 100K splice acceptor site
 ; NAME/KEY: gene
 ; LOCATION: (23680)..(27886)
 ; OTHER INFORMATION: /gene: L4
 ; NAME/KEY: polyA_site
 ; LOCATION: (27920)
 ; NAME/KEY: misc_feature
 ; LOCATION: (28315)
 ; OTHER INFORMATION: /note= fibre splice acceptor site
 ; NAME/KEY: misc_feature
 ; LOCATION: (28341)
 ; OTHER INFORMATION: /note= fibre splice acceptor site
 ; NAME/KEY: gene
 ; LOCATION: (28363)..(31768)
 ; OTHER INFORMATION: /gene: L5
 ; NAME/KEY: misc_feature
 ; LOCATION: (30511)
 ; OTHER INFORMATION: /gene: L5 /note= fibre splice acceptor site
 ; NAME/KEY: polyA_site
 ; LOCATION: (31770)
 ; US-09-970-711-1

Alignment Scores:
 Pred. No.: 2.44e+03 Length: 43804
 Score: 36.00 Matches: 7
 Percent Similarity: 90.00% Conservative: 2
 Best Local Similarity: 70.00% Mismatches: 1
 Query Match: 73.47% Indels: 0
 DB: 10 Gaps: 0
 US-09-823-649A-5 (1-11) x US-09-970-711-1 (1-43804)

OY 2 ServalArgLengthly***Provalysgiu 11
 Db 32613 TCGGTGACGTGAGGAGATCCGGTGAAGAG 32642

RESULT 7

US-10-004-717-18
 ; Sequence 18, Application US/10004717
 ; Publication No. US2002019265A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZOGHEI, HUDA Y.
 ; APPLICANT: YANG, QI
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
 ; TITLE OF INVENTION: ATOMAL ASSOCIATED SEQUENCE FOR DEAFNESS,
 ; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
 ; FILE REFERENCE: P01899US4
 ; CURRENT APPLICATION NUMBER: US/10/004,717
 ; CURRENT FILING DATE: 2002-08-16
 ; PRIOR APPLICATION NUMBER: 09/585,645
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: 60/176,993
 ; PRIOR FILING DATE: 2000-01-19
 ; PRIOR APPLICATION NUMBER: 60/137,060
 ; PRIOR FILING DATE: 1999-06-01
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 1074
 ; TYPE: DNA
 ; ORGANISM: chicken
 ; US-10-004-717-18

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
62	35.00	1074	7	1	2	0	0
80.00%							
Best Local Similarity:	70.00%						
Query Match:	71.43%						

US-09-823-649A-5 (1-11) x US-10-004-717-18 (1-1074)

OY 2 ServalArgLengthly***Provalysgiu 11

Db 972 TCCATGACCTTGAAACCGGCAAGAA 1001

RESULT 8

US-09-731-872-37/c
 ; Sequence 37, Application US/09731872
 ; Patent No. US20020102604A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, Jean Baptiste
 ; APPLICANT: Bouquelerec, Lydie
 ; APPLICANT: Jobert, Severin
 ; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
 ; FILE REFERENCE: 78,US3,REG
 ; CURRENT APPLICATION NUMBER: US/09/731,872
 ; CURRENT FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/169,629
 ; PRIOR FILING DATE: 1999-12-08
 ; PRIOR APPLICATION NUMBER: US 60/187,470
 ; PRIOR FILING DATE: 2000-03-06
 ; NUMBER OF SEQ ID NOS: 482
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 37
 ; LENGTH: 1945
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 60..1682
 ; NAME/KEY: sig_peptide
 ; LOCATION: 60..143

OTHER INFORMATION: Von Heijne matrix
 ; OTHER INFORMATION: score 3.75144398608723
 ; OTHER INFORMATION: seq SGLLQVLFRLIT/TV
 ; US-09-731-872-37

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
203	34.00	1945	7	0	1	0	0
87.50%							
Best Local Similarity:	87.50%						
Query Match:	69.39%						

US-09-823-649A-5 (1-11) x US-09-731-872-37 (1-1945)

OY 1 LeuServalArgLengthly***Pro 8

Db 1675 TTGTGACGTGCTGGGACACCT 1652

RESULT 9

US-09-731-872-41/c
 ; Sequence 41, Application US/09731872
 ; Patent No. US20020102604A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, Jean Baptiste
 ; APPLICANT: Bouquelerec, Lydie
 ; APPLICANT: Jobert, Severin
 ; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
 ; FILE REFERENCE: 78,US3,REG
 ; CURRENT APPLICATION NUMBER: US/09/731,872
 ; CURRENT FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/169,629
 ; PRIOR FILING DATE: 1999-12-08
 ; PRIOR APPLICATION NUMBER: US 60/187,470
 ; PRIOR FILING DATE: 2000-03-06
 ; NUMBER OF SEQ ID NOS: 482
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 41
 ; LENGTH: 1953
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 56..1678
 ; NAME/KEY: sig_peptide
 ; LOCATION: 56..139
 ; OTHER INFORMATION: Von Heijne matrix
 ; OTHER INFORMATION: score 3.75144398608723
 ; OTHER INFORMATION: seq SGLLQVLFRLIT/TV
 ; US-09-731-872-41

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
204	34.00	1953	7	0	1	0	0
87.50%							
Best Local Similarity:	87.50%						
Query Match:	69.39%						

US-09-823-649A-5 (1-11) x US-09-731-872-41 (1-1953)

OY 1 LeuServalArgLengthly***Pro 8

Db 1671 TTGTGACGTGCTGGGACACCT 1648

RESULT 10

US-09-731-872-59/c
 ; Sequence 59, Application US/09731872
 ; Patent No. US20020102604A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, Jean Baptiste
 ; APPLICANT: Bouquelerec, Lydie
 ; APPLICANT: Jobert, Severin


```

; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 59
; LENGTH: 1969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 35..1657
; NAME/KEY: sig_peptide
; LOCATION: 35..118
; OTHER INFORMATION: Von Heljne matrix
; OTHER INFORMATION: score 3.75144398608723
; OTHER INFORMATION: seq SGLLQVLFRLIT/FV
US-09-731-872-59

Alignment Scores:
Pred. No.: 205 Length: 1969
Score: 34.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-5 (1-11) x US*09-731-872-59 (1-1969)
QY 1 LeuSeVaIArgIeudLy***Pro 8
Db 1650 TTGTCAGTGGCTGAGACACCT 1627

RESULT 11
US-09-815-242-9539
; Sequence 9539, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Yrawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9539

; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
US-09-815-242-9415

Alignment Scores:
Pred. No.: 209 Length: 2001
Score: 34.00 Matches: 6
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 54.55% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-5 (1-11) x US-09-815-242-9539 (1-2001)
QY 1 LeuSeVaIArgIeudLy***ProValIySGIu 11
Db 40 TTATCTGTAAAGGTAGCTTACTATTAAGGAA 72

RESULT 12
US-09-815-242-9415
; Sequence 9415, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Yrawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9415
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
US-09-815-242-9415

Alignment Scores:
Pred. No.: 213 Length: 2031
Score: 34.00 Matches: 6
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 54.55% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSLPYEE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: - *
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/gen2.1/USPRO.spool/US0982369/quat.21012003.093151.24616/app.query.fasta.1.1393
-DB=EST -QFMT=fastcap -SUFFIX=trstc -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0
-UNIT5-bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0982369.0931.1.2820.093151.24616 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEBUFFER -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : EST: *
1: em_estha: *
2: em_esthum: *
3: em_estnu: *
4: em_estrov: *
5: em_estro: *
6: em_estro: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	45	90.0	447 13	B1929077 EST548966
2	45	90.0	480 12	BF050758 EST3435916
3	45	90.0	495 13	B1129400 G090P24V
4	45	90.0	507 14	A1895448 EST264891
5	45	90.0	534 14	BQ489763 05-B9425-
6	45	90.0	592 14	BE434433 EST405511
7	45	90.0	629 14	BQ990012 CGF19D10.
8	45	90.0	728 14	BQ856219 BQ826014.
9	45	90.0	768 13	BM411779 B51386106
10	41	82.0	788 12	BQ216020 RST35834
11	40	80.0	438 13	BM482565 535491 MA
12	39	78.0	183 10	AV415207 AV415207
13	39	78.0	411 10	AV428667 AV428667
14	39	78.0	416 12	BF758509 NK4-CT053
15	39	78.0	426 10	AV415722 AV415722
16	39	78.0	650 17	AZ826009 2M0101C05
17	39	78.0	658 14	BQ404238 CA_F006
18	39	78.0	664 14	BQ139058 NF010E06P
19	39	78.0	674 13	B1308144 EST509554
20	39	78.0	742 9	A1498453 A1498453
21	38	76.0	357 10	AM221400 EST297869
22	38	76.0	408 12	BF904943 11.2-MV017
23	38	76.0	500 17	AQ728992 HS_5466-B
24	38	76.0	501 9	A1483959 EST249830
25	38	76.0	548 10	AM930669 EST36512
26	38	76.0	608 13	EM165312 EST568035
27	38	76.0	611 17	AQ550437 RPT1-11-3
28	38	76.0	686 17	BH488006 BOCUR621F
29	38	76.0	700 17	BH503133 BOCUC387F
30	37	74.0	246 12	BF553693 BT-R-BT0-
31	37	74.0	249 9	AV035497 AV035497
32	37	74.0	321 14	BQ084906 1145607 Y
33	37	74.0	568 14	BQ696662 NKPV_043-
34	37	74.0	571 12	BQ834832 353186 MA
35	37	74.0	587 13	BQ315731 BQ315731
36	37	74.0	652 10	AM319554 UM97404.x
37	37	74.0	653 13	B141990 602991072
38	37	74.0	706 13	B1102747 602888324
39	37	74.0	721 13	BQ321189 BQ321189
40	37	74.0	761 13	B1082481 B1082481
41	37	74.0	764 9	AU131236 AU131236
42	37	74.0	868 13	AU136442 AU136442
43	37	74.0	888 13	BF553537 BF553537
44	37	74.0	928 13	B1903867 B1903867
45	37	74.0	941 12	BQ165134 BQ165134

ALIGNMENTS

RESULT 1
LOCUS B1929077
DEFINITION EST548966 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
clone cTOB27L17 5' end, mRNA sequence.
ACCESSION B1929077
VERSION B1929077.1 GI:16241891
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 447)

AUTHORS van der Hoeven, R.S., Bezzzerides, J.L., Karamycheva, S.A., Tsai, J., Uitterlbeck, T., Van Aken, S., Roning, C.M., Niemman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
O. van der Hoeven, Department of Plant Breeding, Wageningen University, 6700 AH Wageningen, The Netherlands

BASE COUNT	118 a	92 c	89 g	148 t
ORIGIN				

Alignment Scores:	
Prod. No.:	52.6
Score:	45.00
Percent Similarity:	90.91%
Best Local Similarity:	72.73%
Query Match:	90.00%
BB:	13
Length:	447
Matches:	8
Conservative:	2
Mismatches:	1
Indels:	0
Gaps:	0

US-09-823-649A-4 (1-11) X BI929077 (1-447)

```

QY      1 LeuSer**GluLeuSerIleProTyrGluGlu 11
          ||||| |||||:::|||||:::
DB      331 CTCTCGGCTGAGCTCTCAGTTCGGTATGAGGAT 363

```

[illegible]

```

/organism="Lycopersicon esculentum"
/cultivar="FA496"
/db_xref="taxon:4081"
/clone="CLEM19022"
/clone_1lb="tomato developing/Immature green fruit"
/tissue_type="fruit"
/dev_stage="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/note="Vector: pBluescript(SKmutadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."
BASE COUNT      116 a      102 c      119 g      143 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	57.8
Score:	45.00
Percent Similarity:	90.91%
Best Local Similarity:	72.73%
Query Match:	90.00%
DB:	12
Gaps:	0
Length:	480
Matches:	8
Conservative:	2
Mismatches:	1
Indels:	0
Gaps:	0

US-09-823-649A-4 (1-11) x BF050758 (1-480)

Oy 1 Leuser***e|uleuserleProtyrGlulu 11
||||| |||||:::|||||:::
Db 261 CTCGGCTAGCTCTCAATCCGATATAGAT 29

US-09-823-649A-4 (1-11) x BF050758 (1-480)

```

Qy 1 Leuser***GluleuserileProTyrgluin 11
      |||||  |||||:::|||||:::
Db 261 CTCTGGGCTGACCTCTCAGTTCCTATGAGGAT 29

```

RESULT 3				
B1129400				
LOCUS	B1129400	495 bp	mRNA	linear
DEFINITION	G090P24Y Populus cambium cDNA library Populus tremula x Populus tremuloides cDNA, mRNA sequence.			

FEATURES	Location/Qualifiers
source	1. .495

BASE COUNT	109 a	146 c	117 g	123 t
ORIGIN	/orgname="Populus tremula & Populus trichocarpa" /db_xref="taxon:47664" /clone_lib="Populus cambium cDNA library" /note="Organ: cambium"			

Alignment Scores:	
Pred. No.:	60.2
Score:	45.00
Percent Similarity:	92.91%
Best Local Similarity:	70.73%
Query Match:	90.00%
DB:	13
Gaps:	0
Length:	499
Matches:	8
Conservative:	2
Mismatches:	1
Indels:	0
Gaps:	0

US-09-823-649a-4 (1-11) x B1129400 (1-495)

OY 1 Leuser***GluuSer11leProtyGluu 11
||||| |||||||:|||||:|||||:|
Db 59 CTCTCGGCGAGTTCGCTGCTCTATGAGAT 91

RESULT 4
LOCUS A1895448 507 bp mRNA linear EST 18-MAY-2001
DEFINITION EST264891 tomato callus, TAMU Lycopersicon esculentum cDNA clone
cLEC7H2, mRNA sequence.
ACCESSION A1895448
VERSION A1895448.1 GI:5601350
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 507)
Alcala,T., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Konning,
C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.*

FEATURES
SOURCE Location/Qualifiers
1..507
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEC7H2"
/clone_lib="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="X11-BIue MRP"
/note="Vector: pBluescript SK(-); Site.1: Kcor1; Site.2:
Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST library"

BASE COUNT 125 a 138 c 114 g 130 t
ORIGIN

Alignment Scores:

Pred. No.: 62.1 Length: 507
Score: 45.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 90.00% Indels: 0
Db: 9 Gaps: 0

US-09-823-649a-4 (1-11) x A1895448 (1-507)

OY 1 Leuser***GluuSer11leProtyGluu 11
||||| |||||||:|||||:|||||:|
Db 66 CTCTCGGCGAGTTCGCTGCTCTATGAGAT 98

RESULT 5
LOCUS BQ489763 534 bp mRNA linear EST 07-JUN-2002
DEFINITION 05-E9425-006-009-101-T3 Sugar beet Mp12-ADIS-006 Lambda zap II
1-Beta vulgaris cDNA clone I-1-9, mRNA sequence.
ACCESSION BQ489763
VERSION BQ489763.1 GI:21334383

KEYWORDS EST.
SOURCE Beta vulgaris.
ORGANISM Beta vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.

REFERENCE 1 (bases 1 to 534)
Bellin,D., Werber,M., Theis,T., Weissshaar,B. and Schneider,K.
EST sequencing, annotation and macroarray expression analysis of
more than 3000 sugar beet cDNAs identifies genes with root-specific
expression pattern
Unpublished (2002)
CONTACT: Weissshaar B
ADIS DNA core facility at Mp12
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weissshaar@mp12-koeln.mpg.de
Seq primer: T3 'AATTAACCCGACCTAAGCG'
High quality sequence stop: 534.

FEATURES
SOURCE Location/Qualifiers
1..534
/organism="Beta vulgaris"
/db_xref="taxon:161934"
/clone="1-1-9"
/clone_lib="Sugar beet Mp12-ADIS-006 Lambda zap II
library"
/dev_stage="4 week old pot-grown plants"
/note="Organ: shoot and root. Vector: pBluescript SK- from
lambda ZAP II: cDNA (lambda ZAP-II) library from sugar
beet, whole plant mRNA. Prepared using the Strategene
UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites
and orientation:
rev-T3-SacI-SK-EcoRI-GGCGACGAGC-5-pr-cDNA-polyA-XhoI-KpnI-T7
-uni"

BASE COUNT 143 a 112 c 127 g 152 t
ORIGIN

Alignment Scores:

Pred. No.: 66.5 Length: 534
Score: 45.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 90.00% Indels: 0
Db: 14 Gaps: 0

US-09-823-649a-4 (1-11) x BQ489763 (1-534)

OY 1 Leuser***GluuSer11leProtyGluu 11
||||| |||||||:|||||:|||||:|
Db 281 CTCTCTGCTGACCTCTGCTCTATGAGGAC 313

RESULT 6

LOCUS BQ434433 592 bp mRNA linear EST 18-MAY-2001
DEFINITION EST405511 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEC17C9, mRNA sequence.
ACCESSION BQ434433
VERSION BQ434433.1 GI:9432276
KEYWORDS EST.

SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 592)
Alcala,T., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Konning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,
S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)

COMMENT

Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

SOURCE

Location/Qualifiers
1..592
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEG17C9"
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmClnadp1; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 140 a 149 c 137 g 166 t
ORIGIN

Alignment Scores:

Pred. No.:	76.1	Length:	592
Score:	45.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	90.00%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649a-4 (1-11) x BB434433 (1-592)

QY 1 Leuser***Glulenser1leprotYrGluc11
||||| |||||||:|||||:|
DB 194 CTCCTGCGTGCCTCAGTCCGTATGAGGAT 226

RESULT 7

LOCUS

BQ990012 629 bp mRNA linear NSI 21-AUG-2002
DEFINITION QGF19D10.y9.ab1 QG_EFGHT lettuce serriola Lactuca sativa cDNA clone

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]belongs to contig QG_CA.Contig1419, see <http://cgpb.ucdavis.edu/>

for details.

Plate: QGF19 row: D column: 10.

Location/Qualifiers

1..629

/organism="Lactuca sativa"

/cultivar="L.serriola"

/db_xref="taxon:4236"

/clone="QGF19D10"

/clone_lib="QG_EFGHU lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBRCDNA5142; The library was constructed from 10 different sources of cDNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpb.ucdavis.edu/>

/tag_lib="QG_EFGHU lettuce serriola"

/tag_libsub="flowers post-fertilized"

/tag_seq="GCCATTCGCG"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	92.4	Length:	629
Score:	45.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	90.00%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649a-4 (1-11) x BQ990012 (1-629)

QY 1 Leuser***Glulenser1leprotYrGluc11
||||| |||||||:|||||:|
DB 471 CTCCTGCGTGCCTCAGTCCGTATGAGGAT 439

RESULT 8

LOCUS

BQ856219 728 bp mRNA linear EST 14-AUG-2002
DEFINITION QGB28014.y9.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]belongs to contig QG_CA.Contig1419, see <http://cgpb.ucdavis.edu/>

for details.

Plate: QGB28 row: O column: 14.

Location/Qualifiers

1..728

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="QGB28014"

/clone_lib="QG_ABCDI lettuce salinas"


```

|||||
Db      546  CTGTCGCAAACTTCATTGCTTACGAGAA 514
|||||

RESULT 11
LOCUS   BM482565/c      438 bp  mRNA  linear  EST 05-FEB-2002
DEFINITION  535491 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BM482565
VERSION    BM482565.1  GI:18532893
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS   Smith,T.P., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahnenkratz,S.C., Bennett,
            G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chilkotkova,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Kele,J.W.
            1 (bases 1 to 438)
            Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
            Genome Res. 11 (4), 626-630 (2001)
            Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and all trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -mismatch 12 options.
            PCR primers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTCCTCATCGACG
            Plate: 3 row: 1 column: 18
            Seq primer: ATTTCGTGACACTATAG.
            Location/Qualifiers
                source
                    1..438
                        /organism="Bos taurus"
                        /db_xref="taxon:9913"
                        /clone_11b="MARC 3BOV"
                        /tissue_type="pooled"
                        /lab_host="DH10B"
                        /note="vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
                        Library made from pooled tissue from marrow, alveolar
                        macrophage, ovary, fetal semitendinosus muscle, and fetal
                        longissimus muscle."
BASE COUNT      126 a      84 c      90 g      138 t
ORIGIN
Alignment Scores:
    Pred. No.:      476      Length:      438
    Score:          40.00      Matches:      7
    Percent Similarity: 100.00%      Conservative: 1
    Best Local Similarity: 87.50%      Mismatches: 0
    Query Match:      80.00%      Indels:      0
    DB:              13      Gaps:      0

US-09-823-649a-4 (1-11) x BM482565 (1-438)

QY      4  GluLeuSerIleProTyrGluGlu 11
        |||||||:|||||
Db      313  GAATATCTGTCTCCCTACGAGAG 290

RESULT 12
LOCUS   AV415207      183 bp  mRNA  linear  EST 23-MAY-2000
DEFINITION  AV415207 Lotus japonicus young plants (two-week old) lotus
            japonicus cDNA clone MM104d04_r_5', mRNA sequence.
ACCESSION  AV415207

```

```

VERSION      AV415207.1  GI:7744383
KEYWORDS     EST.
SOURCE       Lotus japonicus.
ORGANISM     Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
            Lotus.
            1 (bases 1 to 183)
            Asamizu,K., Nakamura,Y., Sato,S. and Tabata,S.
            Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
            DNA Res. 7 (2), 127-130 (2000)
            Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
            Location/Qualifiers
                source
                    1..183
                        /organism="Lotus japonicus"
                        /db_xref="taxon:34105"
                        /clone_11b="MM104d04_r_5'"
                        /dev_stage="young plants (two-week old)"
                        /note="vector: phuscriptplII SK-; Site_1: EcoRI; Site_2:
                        XhoI; isolate=Miyakojima MG-20"
BASE COUNT      33 a      64 c      44 g      42 t
ORIGIN
Alignment Scores:
    Pred. No.:      236      Length:      183
    Score:          39.00      Matches:      7
    Percent Similarity: 81.82%      Conservative: 2
    Best Local Similarity: 63.64%      Mismatches: 2
    Query Match:      78.00%      Indels:      0
    DB:              10      Gaps:      0

US-09-823-649a-4 (1-11) x AV415207 (1-183)

QY      1  LeuSer**GluLeuSerIleProTyrGluGlu 11
        |||||||:|||||
Db      117  CTCTCGCGCGAGTCTGTCTTTCATGATGAGAA 149

RESULT 13
LOCUS   AV428667      411 bp  mRNA  linear  EST 02-MAY-2000
DEFINITION  AV428667 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MM1050g03_r_5', mRNA sequence.
ACCESSION  AV428667
VERSION    AV428667.1  GI:7678049
KEYWORDS   EST.
SOURCE     Lotus japonicus.
ORGANISM   Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
            Lotus.
            1 (bases 1 to 411)
            Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
            Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
            DNA Res. 7 (2), 127-130 (2000)
            Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
            Location/Qualifiers
                source
                    1..411
                        /organism="Lotus japonicus"

```

```

/db_xref="taxon:34305"
/clone="MM1050g03.r"
/clone_lib="Lotus japonicus young plants (two-week old)"
/dev_stage="young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; isolate-MiyakoJima MG-20"
BASE COUNT      86 a      143 c      90 g      90 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      683      Length:      411
Score:          39.00      Matches:      7
Percent Similarity: 81.82%      Conservative: 2
Best Local Similarity: 63.64%      Mismatches: 2
Query Match:    78.00%      Indels:      0
DB:              10      Gaps:        0

US-09-823-649a-4 (1-11) x AV428667 (1-411)

QY      1 Leuser***Gluleuserilleprotyrcluglu 11
        ||||| ||||| ::|||:|||
Db      374 CTCTCCGCGACGCTCTGTCTGCGATGATGAA 406

RESULT 14
BF758509/c      BF758509      416 bp      mRNA      linear      EST 12-JAN-2001
LOCUS      MR4-CT0538-141100-104-D08 CT0538 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF758509
VERSION      BF758509.1 GI:12106409
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 416)
AUTHORS      D. S. Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Naal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR4&ct=MR4-CT0538-
141100-104-D08&ts=2000-11-14&f=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 88.
Location/Qualifiers
FEATURES
source
1. 416
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0538"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

```

```

BASE COUNT      129 a      106 c      85 g      96 t
ORIGIN

Alignment Scores:
Pred. No.:      694      Length:      416
Score:          39.00      Matches:      8
Percent Similarity: 81.82%      Conservative: 1
Best Local Similarity: 72.73%      Mismatches: 2
Query Match:    78.00%      Indels:      0
DB:              12      Gaps:        0

US-09-823-649a-4 (1-11) x BF758509 (1-416)

QY      1 Leuser***Gluleuserilleprotyrcluglu 11
        ||||| ||||| ::|||:|||
Db      182 TTGCTGTGAGACGACTGCTTATGAGAG 150

RESULT 15
AV415722
LOCUS      AV415722 Lotus japonicus young plants (two-week old) Lotus
DEFINITION      japonicus cDNA clone MM115e12_r 5', mRNA sequence.
ACCESSION      AV415722
VERSION      AV415722.1 GI:7744898
KEYWORDS      EST.
SOURCE      Lotus japonicus.
ORGANISM      Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE      1 (bases 1 to 426)
AUTHORS      Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE      Generation of 7137 non-redundant expressed sequence tags from a
legume, Lotus japonicus
JOURNAL      DNA Res. 7 (2), 127-130 (2000)
MEDLINE      20277479
COMMENT      Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yata 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamura@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
FEATURES
source
1. 426
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone="MM115e12.r"
/clone_lib="Lotus japonicus young plants (two-week old)"
/dev_stage="Young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; isolate-MiyakoJima MG-20"
BASE COUNT      90 a      147 c      95 g      94 t
ORIGIN

Alignment Scores:
Pred. No.:      716      Length:      426
Score:          39.00      Matches:      7
Percent Similarity: 81.82%      Conservative: 2
Best Local Similarity: 63.64%      Mismatches: 2
Query Match:    78.00%      Indels:      0
DB:              10      Gaps:        0

US-09-823-649a-4 (1-11) x AV415722 (1-426)

QY      1 Leuser***Gluleuserilleprotyrcluglu 11
        ||||| ||||| ::|||:|||
Db      375 CTCTCCGCGACGCTGTGTTCCGTATGATGAA 407

Search completed: January 21, 2003, 12:24:18
Job time : 1170.14 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame-plus-p2n model

Run on: January 21, 2003, 09:31:57 : Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649a-4
Perfect score: 50
Sequence: 1 LKXELIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 393868 segs, 222934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODE=frame-p2n model -DEV=xlid
-Q/cgnt2.1/USPRO.spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published Applications_NA -QPM=fastlap -SUFFIX=trnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62
-TRANS=human4.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09823649_@CGN 1.1.77 @runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO.XLPHY -NO.MMAP -LARGEBUQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*
1: /cgnt2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgnt2_6/ptodata/1/pubpna/PCNT_NEW_PUB.seq:*
3: /cgnt2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgnt2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgnt2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgnt2_6/ptodata/1/pubpna/PCNTUS_PUBCOMB.seq:*
7: /cgnt2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgnt2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgnt2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgnt2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
11: /cgnt2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
12: /cgnt2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
13: /cgnt2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgnt2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	96.0	2496	9	US-10-033-297-2
2	48	96.0	2496	9	US-10-081-806-2
3	45	90.0	1600	9	US-10-033-297-12
4	45	90.0	1600	9	US-10-081-806-12

5	45	90.0	2502	9	US-10-033-297-7	Sequence 7, Appl1
6	45	90.0	2502	9	US-10-033-297-21	Sequence 21, Appl1
7	45	90.0	2502	9	US-10-081-806-7	Sequence 7, Appl1
8	45	90.0	2502	9	US-10-081-806-21	Sequence 21, Appl1
9	45	90.0	2504	9	US-10-033-297-3	Sequence 3, Appl1
10	45	90.0	2504	9	US-10-081-806-3	Sequence 3, Appl1
11	45	90.0	2505	9	US-10-033-297-65	Sequence 65, Appl1
12	45	90.0	2505	9	US-10-033-297-68	Sequence 68, Appl1
13	45	90.0	2505	9	US-10-033-297-70	Sequence 70, Appl1
14	45	90.0	2506	9	US-10-033-297-1	Sequence 1, Appl1
15	45	90.0	2506	9	US-10-081-806-1	Sequence 1, Appl1
16	45	90.0	2511	10	US-09-777-430A-9	Sequence 9, Appl1
17	45	90.0	2511	10	US-09-777-430A-12	Sequence 12, Appl1
18	45	90.0	2526	10	US-09-777-430A-14	Sequence 14, Appl1
19	45	90.0	2526	10	US-09-777-430A-15	Sequence 15, Appl1
20	45	90.0	2526	10	US-09-777-430A-25	Sequence 25, Appl1
21	45	90.0	2526	10	US-09-777-430A-25	Sequence 25, Appl1
22	45	90.0	2526	10	US-09-777-430A-25	Sequence 25, Appl1
23	37	74.0	1497	9	US-09-725-169A-3	Sequence 3, Appl1
24	37	74.0	1497	10	US-09-735-171A-3	Sequence 3, Appl1
25	37	74.0	1497	10	US-09-735-171A-3	Sequence 3, Appl1
26	37	74.0	1497	10	US-09-735-171A-3	Sequence 3, Appl1
27	36	72.0	286	10	US-09-923-876-6139	Sequence 6139, Ap
28	36	72.0	2603	9	US-09-764-868-80	Sequence 80, Appl1
29	34	68.0	302	9	US-10-046-935-1913	Sequence 1913, Ap
30	34	68.0	302	9	US-09-878-178-1913	Sequence 1913, Ap
31	34	68.0	378	10	US-09-734-017A-57	Sequence 57, Appl1
32	34	68.0	402	10	US-09-974-300-3201	Sequence 3201, Ap
33	34	68.0	675	10	US-09-841-132-377	Sequence 377, App
34	34	68.0	675	10	US-09-841-132-473	Sequence 473, App
35	34	68.0	1399	10	US-09-764-877-134	Sequence 134, App
36	34	68.0	1812	10	US-09-925-300-724	Sequence 724, App
37	34	68.0	2068	12	US-10-044-090-519	Sequence 519, App
38	34	68.0	2515	10	US-10-044-090-520	Sequence 520, App
39	34	68.0	3664	10	US-09-764-877-2933	Sequence 2933, Ap
40	34	68.0	12445	10	US-09-070-927A-242	Sequence 242, App
41	34	68.0	13293	10	US-09-764-877-3404	Sequence 3404, Ap
42	34	68.0	16281	10	US-09-764-847-1367	Sequence 1367, Ap
43	34	68.0	16285	10	US-09-764-847-1368	Sequence 1368, Ap
44	34	68.0	16285	10	US-09-764-847-1369	Sequence 1369, Ap
45	34	68.0	30515	10	US-09-764-847-1208	Sequence 1208, Ap

ALIGNMENTS

RESULT 1
US-10-033-297-2
Sequence 2, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Mary Ann L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESS: Medien & Carroll, LLP
SERRET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NO. US20020187486A1-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-033-297-2

Alignment Scores:
Pred. No.: 0.501 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-4 (1-11) x US-10-033-297-2 (1-2496)
QY 1 Leuser***Gluleuser1leProtyrGluGlu 11
||||| |||||||
Db 2029 CTCCTCGGAGAGCTTCATCCCTACAGAGAG 2061

RESULT 2
US-10-081-806-2
Sequence 2, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-081-806-2

Alignment Scores:
Pred. No.: 0.501 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 9 Caps: 0

US-09-823-649a-4 (1-11) x US-10-081-806-2 (1-2496)
QY 1 Leuser***Gluleuser1leProtyrGluGlu 11
||||| |||||||
Db 2029 CTCCTCGGAGAGCTTCATCCCTACAGAGAG 2061

RESULT 3
US-10-033-297-12
Sequence 12, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-No. US20020187486A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072

FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-033-297-12
Alignment Scores:
Pred. No.: 1,32 Length: 1600
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-033-297-12 (1-1600)
QY 1 Leuser**GluLeuserIleProTyrGluGlu 11
Db 1133 CTCCTCCAGAGACTAGCCATCCCTTACGAGAG 1165
RESULT 4
US-10-081-806-12
Sequence 12, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996

APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-081-806-12
Alignment Scores:
Pred. No.: 1,32 Length: 1600
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-081-806-12 (1-1600)
QY 1 Leuser**GluLeuserIleProTyrGluGlu 11
Db 1133 CTCCTCCAGAGACTAGCCATCCCTTACGAGAG 1165
RESULT 5
US-10-033-297-7
Sequence 7, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Mast, Andrea L.
Lyamichev, Victor I.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 12-DEC-1996
APPLICATION NUMBER: US 08/756,386

FILED DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-033-297-7
Alignment Scores:
Pred. No.: 2.19 length: 2502
Score: 45.00 matches: 9
Percent Similarity: 90.91% conservative: 1
Best Local Similarity: 81.82% mismatches: 1
Query Match: 90.00% indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-033-297-7 (1-2502)
QY 1 Leuser***GluserMeproTyrGluGlu 11
||||| |||||:|||||
Db 2035 CTCTCCAGGAGCTTGCCATCCCTTACGAGGAG 2067
RESULT 5
US-10-033-297-21
Sequence 21, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996

APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-033-297-21
Alignment Scores:
Pred. No.: 2.19 length: 2502
Score: 45.00 matches: 9
Percent Similarity: 90.91% conservative: 1
Best Local Similarity: 81.82% mismatches: 1
Query Match: 90.00% indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-033-297-21 (1-2502)
QY 1 Leuser***GluserMeproTyrGluGlu 11
||||| |||||:|||||
Db 2035 CTCTCCAGGAGCTTGCCATCCCTTACGAGGAG 2067
RESULT 7
US-10-081-806-7
Sequence 7, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-FEB-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-081-806-7
Alignment Scores:
Pred. No.: 2.19 Length: 2502
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-081-806-7 (1-2502)
QY 1 LeuSer**GluLeuSerIleProTyrGlu 11
||||| |||||:|||||||
Db 2035 CTCCTCCAGAGGCTGCATCCCTACGAGAG 2067
RESULT 8
US-10-081-806-21
Sequence 21, Application US9/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
Street: 220 Montgomery Street, Suite 2200
City: San Francisco
State: California
Country: United States Of America
Zip: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-081-806-21
Alignment Scores:
Pred. No.: 2.19 Length: 2502
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-081-806-21 (1-2502)
QY 1 LeuSer**GluLeuSerIleProTyrGlu 11
||||| |||||:|||||||
Db 2035 CTCCTCCAGAGGCTGCATCCCTACGAGAG 2067
RESULT 9
US-10-033-297-3
Sequence 3, Application US/1004329/
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
Street: 220 Montgomery Street, Suite 2200
City: San Francisco
State: California
Country: United States Of America
Zip: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-Nov-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/99/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2504 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-033-297-3

Alignment Scores:
Pred. No.: 2.19 Length: 2504
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-4 (1-11) x US-10-033-297-3 (1-2504)
OY 1 1 User**GlucoseSerileProTyGluGlu 11
||||| |||||:||||| |||||
Db 2038 CTCCTCCAGAGCTTCCCATCCCTACGAGAG 2070

RESULT 10
US-10-081-806-3
Sequence 3, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2504 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-081-806-3

Alignment Scores:
Pred. No.: 2.19 Length: 2504
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-4 (1-11) x US-10-081-806-3 (1-2504)
OY 1 1 User**GlucoseSerileProTyGluGlu 11
||||| |||||:||||| |||||
Db 2038 CTCCTCCAGAGCTTCCCATCCCTACGAGAG 2070

RESULT 11
US-10-033-297-65
Sequence 65, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brown, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-033-297-65

Alignment Scores:
Pred. No.: 2.19 Length: 2505
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.828 Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-4 (1-11) x US-10-033-297-65 (1-2505)

QY 1 LeuSer***GlueuSer11eProTyrglu1u 11
||||| ||||||:|||||||||||
DB 2035 CTCCTCCAGAGAGCTAGCCATCCCTTACGAGAG 2067

RESULT 12
US-10-033-297-68
Sequence 68, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"

FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-033-297-68

Alignment Scores:
Pred. No.: 2.19 Length: 2505
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.828 Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-4 (1-11) x US-10-033-297-68 (1-2505)

QY 1 LeuSer***GlueuSer11eProTyrglu1u 11
||||| ||||||:|||||||||||
DB 2035 CTCCTCCAGAGAGCTAGCCATCCCTTACGAGAG 2067

RESULT 13
US-10-033-297-70
Sequence 70, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-033-297-70
Alignment Scores:
Pred. No.: 2.19 length: 2505
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-033-297-70 (1-2505)
QY 1 Leuser**GluLeuserIleProTyGluGlu 11
DB 2035 CTCCTCCAGGAGCTACGACCTTCACGAGAG 2067
RESULT 14
US-10-033-297-1
Sequence 1, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Brow, Mary Ann D.
Mast, Andrea L.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314

FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-033-297-1
Alignment Scores:
Pred. No.: 2.19 length: 2506
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-4 (1-11) x US-10-033-297-1 (1-2506)
QY 1 Leuser**GluLeuserIleProTyGluGlu 11
DB 2032 CTCCTCCAGGAGCTACGACCTTCACGAGAG 2064
RESULT 15
US-10-081-806-1
Sequence 1, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: F0RS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-081-806-1

Alignment Scores:
Pred. No.: 2.19 Length: 2506
Score: 45.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-4 (1-11) x US-10-081-806-1 (1-2506)

QY 1 LeuSer**GlulLeuSerIleProTyrGluGlu 11
||||| |||||:|||||
Db 2032 CTCTCCAGAGACTGACATCCCTTACGAGAG 2064

Search completed: January 21, 2003, 10:07:46
Job time : 31.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 : Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649a-3
Perfect score: 54
Sequence: 1 LSOELAIPTRE 11

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -
-MODE=frame_plus_p2n_model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=published_applications_NA -OPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPT=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTEXT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0
-MATE=200000000 -USER=US09823649_@cgn_1_1_77-@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -IARGUMENTS -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEROUT=120 -WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6
-YGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	54	100.0	9	US-10-033-297-12
2	54	100.0	9	US-10-081-806-21
3	54	100.0	9	US-10-033-297-7
4	54	100.0	9	US-10-033-297-21

5	54	100.0	9	US-10-081-806-7	Sequence 7, Appl1
6	54	100.0	9	US-10-081-806-21	Sequence 21, Appl1
7	54	100.0	9	US-10-033-297-3	Sequence 3, Appl1
8	54	100.0	9	US-10-081-806-3	Sequence 3, Appl1
9	54	100.0	9	US-10-033-297-65	Sequence 65, Appl1
10	54	100.0	9	US-10-033-297-68	Sequence 68, Appl1
11	54	100.0	9	US-10-033-297-70	Sequence 70, Appl1
12	54	100.0	9	US-10-033-297-1	Sequence 1, Appl1
13	54	100.0	9	US-10-081-806-1	Sequence 1, Appl1
14	54	100.0	9	US-09-777-430A-9	Sequence 12, Appl1
15	54	100.0	9	US-09-777-430A-12	Sequence 14, Appl1
16	54	100.0	9	US-09-777-430A-14	Sequence 19, Appl1
17	54	100.0	9	US-09-777-430A-19	Sequence 22, Appl1
18	54	100.0	9	US-09-777-430A-22	Sequence 25, Appl1
19	54	100.0	9	US-09-777-430A-25	Sequence 1, Appl1
20	54	100.0	9	US-09-972-834-1	Sequence 2, Appl1
21	44	81.5	9	US-10-033-297-2	Sequence 2, Appl1
22	44	81.5	9	US-10-081-806-2	Sequence 9, Appl1
23	40	74.1	9	US-09-972-834-9	Sequence 80, Appl1
24	39	72.2	9	US-09-764-868-80	Sequence 128, Appl1
25	39	72.2	9	US-09-070-927A-128	Sequence 6139, Appl1
26	38	70.4	9	US-09-923-876-6139	Sequence 818, Appl1
27	37	68.5	9	US-09-920-300A-818	Sequence 818, Appl1
28	37	68.5	9	US-10-033-297-10	Sequence 10, Appl1
29	37	68.5	9	US-09-925-300-724	Sequence 724, Appl1
30	36	66.7	9	US-09-881-752A-301	Sequence 301, Appl1
31	36	66.7	9	US-09-895-913A-141	Sequence 141, Appl1
32	36	66.7	9	US-09-815-242-6393	Sequence 6393, Appl1
33	36	66.7	9	US-10-033-297-10	Sequence 10, Appl1
34	36	66.7	9	US-10-081-806-10	Sequence 290, Appl1
35	36	66.7	9	US-09-070-927A-290	Sequence 1034, Appl1
36	35	64.8	9	US-09-815-242-1054	Sequence 1054, Appl1
37	35	64.8	9	US-09-960-352-1424	Sequence 4, Appl1
38	35	64.8	9	US-09-816-127-4	Sequence 3918, Appl1
39	35	64.8	9	US-09-815-242-3918	Sequence 6483, Appl1
40	35	64.8	9	US-09-815-242-6483	Sequence 53, Appl1
41	35	64.8	9	US-10-114-883-53	Sequence 3991, Appl1
42	35	64.8	9	US-09-764-877-3991	Sequence 49, Appl1
43	35	64.8	9	US-09-739-254-49	Sequence 49, Appl1
44	35	64.8	9	US-09-904-615-49	Sequence 228, Appl1
45	35	64.8	9	US-09-070-927A-228	

ALIGNMENTS

RESULT 1
US-10-033-297-12
: Sequence 12, Application US/10033297
: Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff C.
Mast, Andrea J.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids by Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NO. US20020187486A1-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1998
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-033-297-12

Alignment Scores:
Pred. No.: 0.0165 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-3 (1-11) x US-10-033-297-12 (1-1600)

QY 1 Leuserglnleuvalaileprotyrghlu 11
DB 1133 CTCTCCAGAGGCTAGCCATCCCTTACGAGAG 1165

RESULT 2
US-10-081-806-12
Sequence 12, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-081-806-12

Alignment Scores:
Pred. No.: 0.0165 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-3 (1-11) x US-10-081-806-12 (1-1600)

QY 1 Leuserglnleuvalaileprotyrghlu 11
DB 1133 CTCTCCAGAGGCTAGCCATCCCTTACGAGAG 1165

RESULT 3
US-10-033-297-7
Sequence 7, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Brow, Mary Ann D.
Masc, Andrea L.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequenced Invasive Cleavages
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1998
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072

FILED DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-033-297-7
Alignment Scores:
Pred. No.: 0.0276 Length: 2502
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-3 (1-11) x US-10-033-297-7 (1-2502)
OY 1 LeuSerGlnGluAlaIleProTYrGluGlu 11
Db 2035 CTCCTCCAGAGACTGTCATCCCTACGAGAG 2067
RESULT 4
US-10-033-297-21
Sequence 21, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids by Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999

APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-033-297-21
Alignment Scores:
Pred. No.: 0.0276 Length: 2502
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-3 (1-11) x US-10-033-297-21 (1-2502)
OY 1 LeuSerGlnGluAlaIleProTYrGluGlu 11
Db 2035 CTCCTCCAGAGACTGTCATCCCTACGAGAG 2067
RESULT 5
US-10-081-806-7
Sequence 7, Application US/10081806
Publication No. US20020197626A1
GENERAL INFORMATION:
APPLICANT: Patent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-FEB-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386

FILED DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: F0RS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-081-806-7

Alignment Scores:
Pred. No.: 0.0276 Length: 2502
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x US-10-081-806-7 (1-2502)

QY 1 LeuSerGInGluLeuAla11ProtyrGlu11
DB 2035 C1CTCCGAGGAGCTTGCATCCCTACGAGAG 2067

RESULT 6
US-10-081-806-21
Sequence 21, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: F0RS-02564

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-081-806-21

Alignment Scores:
Pred. No.: 0.0276 Length: 2502
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x US-10-081-806-21 (1-2502)

QY 1 LeuSerGInGluLeuAla11ProtyrGlu11
DB 2035 C1CTCCGAGGAGCTTGCATCCCTACGAGAG 2067

RESULT 7
US-10-033-297-3
Sequence 3, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2504 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-033-297-3

Alignment Scores:
Pred. No.: 0.0277 Length: 2504
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x US-10-033-297-3 (1-2504)

OY 1 LeuserGIngluAlaIleProTyrgluGlu 11
Db 2038 CTCTCCAGAGACTTGCATCCCTACGAGAG 2070

RESULT 8
US-10-081-806-3
Sequence 3, Application US710081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2504 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-081-806-3

Alignment Scores:
Pred. No.: 0.0277 Length: 2504
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x US-10-081-806-3 (1-2504)

OY 1 LeuserGIngluAlaIleProTyrgluGlu 11
Db 2038 CTCTCCAGAGACTTGCATCCCTACGAGAG 2070

RESULT 9
US-10-033-297-65
Sequence 65, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-No. US20020187486A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-033-297-65

Alignment Scores:
Pred. No.: 0.0277 length: 2505
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-3 (1-11) x US-10-033-297-65 (1-2505)
QY 1 LeuserGIngluLeuAlaIleProTyrgluGlu 11
Db 2035 CTCTCCAGAGAGCTAGCCATCCCTTAGAGAGAG 2067

RESULT 10
US-10-033-297-68
Sequence 68, Application US/10033297
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-033-297-68

Alignment Scores:
Pred. No.: 0.0277 length: 2505
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-3 (1-11) x US-10-033-297-68 (1-2505)
QY 1 LeuserGIngluLeuAlaIleProTyrgluGlu 11
Db 2035 CTCTCCAGAGAGCTAGCCATCCCTTAGAGAGAG 2067

RESULT 11
US-10-033-297-70
Sequence 70, Application US/10033297
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 24-JAN-1996

APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
FEATURE:
DESCRIPTION: /desc = "DNA"
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-033-297-70
Alignment Scores:
Pred. No.: 0.0277 Length: 2505
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-3 (1-11) x US-10-033-297-70 (1-2505)
QY 1 LeuserGlncluleuAlaIleProTyrGluGlu 11
Db 2035 CTCCTCCAGGAGCTACCATCCCTTACGAGAG 2067
RESULT 12
US-10-033-297-1
Sequence 1, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516

FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-033-297-1
Alignment Scores:
Pred. No.: 0.0277 Length: 2506
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-3 (1-11) x US-10-033-297-1 (1-2506)
QY 1 LeuserGlncluleuAlaIleProTyrGluGlu 11
Db 2032 CTCCTCCAGGAGCTACCATCCCTTACGAGAG 2064
RESULT 13
US-10-081-806-1
Sequence 1, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prodent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-FEB-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-081-106-1

Alignment Scores:
Pred. No.: 0.0277 Length: 2506
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x US-10-081-806-1 (1-2506)

QY 1 LeuSerGlnGluLeuAlaIleProTYrGluGlu 11

Db 2032 CTCCTCCAGACGTACGATCCCTTACGAGGAG 2064

RESULT 14
US-09-777-430A-9
Sequence 9, Application US/09777430A
Patent No. US20020128465A1
GENERAL INFORMATION:
APPLICANT: Lyamichev, Victor
APPLICANT: Skrzypczynski, Zbigniew
APPLICANT: Allawi, Halim T.
APPLICANT: Wayland, Sarah R.
APPLICANT: Yakova, Tsetska
APPLICANT: Neif, Bruce P.
TITLE OF INVENTION: Charge tags and the Separation of Nucleic Acid Molecules
FILE REFERENCE: FORS-04912
CURRENT APPLICATION NUMBER: US/09/777,430A
CURRENT FILING DATE: 2001-02-06
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 2511
TYPE: DNA
ORGANISM: Thermus thermophilus
US-09-777-430A-9

Alignment Scores:
Pred. No.: 0.0278 Length: 2511
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-3 (1-11) x US-09-777-430A-9 (1-2511)

QY 1 LeuSerGlnGluLeuAlaIleProTYrGluGlu 11

Db 2044 CTCCTCCAGACGTACGATCCCTTACGAGGAG 2076

RESULT 15

US-09-777-430A-12
Sequence 12, Application US/09777430A
Patent No. US20020128465A1
GENERAL INFORMATION:
APPLICANT: Lyamichev, Victor
APPLICANT: Skrzypczynski, Zbigniew
APPLICANT: Allawi, Halim T.
APPLICANT: Wayland, Sarah R.
APPLICANT: Yakova, Tsetska
APPLICANT: Neif, Bruce P.
TITLE OF INVENTION: Charge tags and the Separation of Nucleic Acid Molecules
FILE REFERENCE: FORS-04912
CURRENT APPLICATION NUMBER: US/09/777,430A
CURRENT FILING DATE: 2001-02-06
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 2511
TYPE: DNA
ORGANISM: Thermus thermophilus
US-09-777-430A-12

Alignment Scores:
Pred. No.: 0.0278 Length: 2511
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-3 (1-11) x US-09-777-430A-12 (1-2511)

QY 1 LeuSerGlnGluLeuAlaIleProTYrGluGlu 11

Db 2044 CTCCTCCAGACGTACGATCCCTTACGAGGAG 2076

Search completed: January 21, 2003, 10:07:44
Job time : 32.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 [Search time 1168.14 Seconds

(without alignments) 152.507 Million cell updates/sec

Title: US-09-823-649a-3

Perfect score: 54

Sequence: 1 LSCRLAIPYEE 11

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delep 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: --
-O/cn2.1/USPNC.spool/US09823649/runat.21012003.093151.24616/app.query.fasta.1.1393
-DB-EST -QPM-fastest -SUFFIX=st -MINMATCH=0.1 -LOOPCL=0 -IOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=DIOSUM62 -PRAS=human40.cbl -LST=45
-DOCLIGN=200 -THR-SCORE=pccl -THR-MAX=100 -THR-MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649 -CGN 1.1.2820 -runat.21012003.093151.24616 -MCP=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGOUDERY -NEG-SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
C 1	44	81.5	788	12	BG216020
C 2	42	77.8	319	13	B1921437
C 3	42	77.8	467	12	BE050176
C 4	42	77.8	526	10	AW932704
C 5	42	77.8	528	13	H1932551
C 6	42	77.8	529	10	AM031837
C 7	42	77.8	537	10	AM399580
C 8	42	77.8	543	10	AM930945
C 9	42	77.8	558	10	AM931411
C 10	42	77.8	590	10	AM931055
C 11	42	77.8	622	10	AM737720
C 12	42	77.8	653	10	AM036109
C 13	42	77.8	743	13	B1938283
C 14	42	77.8	759	12	B6124823
C 15	41	75.9	198	9	AV043104
C 16	41	75.9	416	12	BE758509
C 17	41	75.9	447	13	B1929077
C 18	41	75.9	480	12	BE050758
C 19	41	75.9	495	13	B1129400
C 20	41	75.9	507	9	A1895448
C 21	41	75.9	534	14	B0489763
C 22	41	75.9	592	10	BE434433
C 23	41	75.9	629	14	B0990012
C 24	41	75.9	712	17	BE702922
C 25	41	75.9	728	13	B0856219
C 26	41	75.9	768	14	B0411779
C 27	41	75.9	830	17	BH681916
C 28	40	74.1	249	9	AV035497
C 29	40	74.1	388	9	A1356388
C 30	40	74.1	451	10	AM644700
C 31	40	74.1	491	12	BE325103
C 32	40	74.1	583	12	BG456033
C 33	40	74.1	752	17	BH353192
C 34	40	74.1	758	17	A2191827
C 35	39	72.2	183	10	AV415207
C 36	39	72.2	201	14	D62821
C 37	39	72.2	231	9	AV036702
C 38	39	72.2	270	9	AV037125
C 39	39	72.2	274	9	AV034809
C 40	39	72.2	279	9	AV035894
C 41	39	72.2	281	9	AV037570
C 42	39	72.2	284	9	AV035331
C 43	39	72.2	284	10	BE523442
C 44	39	72.2	298	9	AV034428
C 45	39	72.2	306	9	AV225900

ALIGNMENTS

RESULT 1
BG216020/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 788)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith

```

FEATURES      location/Qualifiers
source        1. .319
              /organism="Lycopodium esculentum"

```

BASE COUNT	149 a	72 c	139 g	107 t
------------	-------	------	-------	-------

1 LeuSerGlnIleuAlaIleProTyrGlnIu 11

```

VERSION  AM05103/.1  GL:3890293
KEYWORDS  EST,
SOURCE    +

```

ORGANISM Lycopersicon esculentum
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids I: Solanales: Solanaceae: Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 529)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Hansen, T.S., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S., Rønning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J. Generation of ESTs from tomato callus tissue Unpublished (1999)

TITLE Unpublished (1999)
JOURNAL Contact: CUGI
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..529
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CL6C39P7"
/clone_1lb="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni laboratory; cDNA - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 169 a 93 c 139 g 128 t
ORIGIN

Alignment Scores:
Pred. No.: 181 Length: 529
Score: 42.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 77.78% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-3 (1-11) x AW031837 (1-529)

OY 1 LeuSerGInGluLeuAlaIleProGluGlu 11
|||||
Db 156 TTGTCCGACGACACGATCCCATATCAAG 124

RESULT 7
AM399580 537 bp mRNA linear EST 18-MAY-2001
LOCUS EST310080 L. pennellii trichome, Cornell University Lycopersicon
DEFINITION pennellii cDNA clone cLEP8E23 5', mRNA sequence.
ACCESSION AM399580
VERSION AM399580.1 GI:6918050
KEYWORDS EST.
SOURCE Lycopersicon pennellii.
ORGANISM Lycopersicon pennellii.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids I: Solanales: Solanaceae: Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 537)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Lakey, J., Holt, I.E., Liang, F., Hansen, T.S., Upton, J., Rønning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J. Generation of ESTs from wild tomato (Lycopersicon pennellii) trichomes Unpublished (1999)

TITLE Unpublished (1999)
JOURNAL Contact: CUGI

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..537
/organism="Lycopersicon pennellii"
/db_xref="taxon:28526"
/clone="cLEP8E23"
/clone_1lb="L. pennellii trichome, Cornell University"
/tissue_type="trichome"
/dev_stage="mixed stages"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells highly enriched for trichomes, with minor contamination by other types of leaf cells."

BASE COUNT 163 a 104 c 139 g 131 t
ORIGIN

Alignment Scores:
Pred. No.: 185 Length: 537
Score: 42.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 77.78% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-3 (1-11) x AW399580 (1-537)

OY 1 LeuSerGInGluLeuAlaIleProGluGlu 11
|||||
Db 49 TTGTCCGACGACACGATCCCATATCAAG 17

RESULT 8
AM39945 543 bp mRNA linear EST 18-MAY-2001
LOCUS EST356788 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEP42021 5', mRNA sequence.
ACCESSION AM39945
VERSION AM39945.1 GI:8106346
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids I: Solanales: Solanaceae: Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 543)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F., Upton, J., Hansen, T.S., Craven, M.B., Bowman, C.L., Ahn, S., Rønning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J. Generation of ESTs from tomato fruit tissue Unpublished (1999)

TITLE Unpublished (1999)
JOURNAL Contact: CUGI
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..543
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEP42021"
/clone_1lb="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

VERSION AM737720.1 GI:7646665
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 622)
 van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang,
 R., Hansen, T.S., Craven, M.B., Bowman, C.L., Konning, C.M., Niernan,
 W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
 Generation of ESTs from tomato flower tissue, anthesis
 Unpublished (1999)
 TITLE CUGI
 JOURNAL
 COMMENT
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 FEATURES
 source
 1..622
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CTOD49"
 /clone_lib="tomato flower buds, anthesis, Cornell
 University"
 /tissue_type="flower"
 /dev_stage="anthesis"
 /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; supplier: Tanksley; Flower buds and flowers were
 taken from greenhouse plants (4-8 wks old, TA496). They
 were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."
 BASE COUNT 191 a 116 c 158 g 157 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 227 Length: 622
 Score: 42.00 Matches: 8
 Percent Similarity: 90.91% Conservative: 2
 Best Local Similarity: 72.73% Mismatches: 1
 Query Match: 77.78% Indels: 0
 DB: 10 Gaps: 0
 US-09-823-649a-3 (1-11) x AM737720 (1-622)
 QY 1 LeuSerGInGluLeuAlaIleProTyrGluCtu 11
 |||||
 Db 154 TTGTCCAGCAGCAACAGCAATCCCATATCAAGAG 122
 RESULT 12
 LOCUS AM036109 653 bp mRNA linear EST 18-MAY-2001
 DEFINITION EST274485 tomato seed, TMMU Lycopersicon esculentum cDNA clone
 CLEB1M9 similar to ATP-dependent protease (CD4b), mRNA sequence.
 ACCESSION AM036109
 VERSION AM036109.1 GI:5894788
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 653)
 Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upton, J., Konning, C.M., Craven, M.B., Fujii, C.Y., Bowman,
 C.L., Niernan, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley,
 S.D. and Giovannoni, J.J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT

COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 3 prime sequence.
 FEATURES
 source
 1..653
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1M9"
 /clone_lib="tomato seed, TMMU"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="X11-Blue MRP"
 /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; clone - tomato seed EST library, directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."
 BASE COUNT 227 a 137 c 105 g 184 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 243 Length: 653
 Score: 42.00 Matches: 8
 Percent Similarity: 90.91% Conservative: 2
 Best Local Similarity: 72.73% Mismatches: 1
 Query Match: 77.78% Indels: 0
 DB: 10 Gaps: 0
 US-09-823-649a-3 (1-11) x AM036109 (1-653)
 QY 1 LeuSerGInGluLeuAlaIleProTyrGluCtu 11
 |||||
 Db 160 TTGTCCAGCAGCAACAGCAATCCCATATCAAGAG 192
 RESULT 13
 LOCUS B1933283/C 743 bp mRNA linear EST 18-OCT-2001
 DEFINITION EST551172 tomato flower, anthesis Lycopersicon esculentum cDNA
 clone cTOD16E14 5' end, mRNA sequence.
 ACCESSION B1933283
 VERSION B1933283.1 GI:16247755
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 743)
 utterback, T., Van Aken, S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,
 Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
 Generation of ESTs from tomato flower tissue, anthesis (2001)
 Unpublished (2001)
 TITLE
 JOURNAL
 COMMENT
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3.
 FEATURES
 source
 1..743
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CTOD16E14"
 /clone_lib="tomato flower, anthesis"
 /tissue_type="flower"

```

/dev_stage="anthesis"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; flower buds and flowers
were taken from greenhouse plants (4-8 wks old, T9496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
BASE COUNT      216 a      135 c      218 g      174 t
ORIGIN
Alignment Scores:
Pred. No.:      290      Length:      743
Score:          42.00     Matches:      8
Percent Similarity: 90.91%  Conservative: 2
Best Local Similarity: 77.73%  Mismatches:  1
Query Match:     77.78%     Indels:      0
DB:              13       Gaps:        0

US-09-823-649a-3 (1-11) x B1933283 (1-743)
QY      1 leuserglnluleaialleprotyrglugu 11
|||||.....|
Db      509 TTGTCCACACACACATCCATTCACAGG 477

RESULT 14
BG124823      759 bp  mRNA  linear  EST 31-JAN-2001
LOCUS      ES1470469 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION  cT066H13 5' sequence, mRNA sequence.
ACCESSION  BG124823
VERSION    BG124823.1  GI:12625011
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; euclcotyledons; core eudicots;
            Asteridae; easterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE  1 (bases 1 to 759)
AUTHORS   van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,
            Hansen, C., Renning, C. and Tanksley, S.
            Generation of ESTs from tomato shoot/meristem tissue
            Unpublished (2001)
JOURNAL
COMMENT    Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
            Location/Qualifiers
                source
                1..759
                /organism="Lycopersicon esculentum"
                /cultivar="T496"
                /db_xref="taxon:4081"
                /clone="cT066H13"
                /clone_lib="tomato shoot/meristem"
                /tissue_type="shoot/meristem"
                /dev_stage="developing shoots from 4-6wks old plants"
                /lab_host="SOLR"
                /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                XhoI; Small expanding leaves from the growing tip were
                taken from greenhouse plants (4-6wks old T9496). Tissue
                was immediately frozen in liquid nitrogen."
BASE COUNT      250 a      140 c      198 g      171 t
ORIGIN
Alignment Scores:
Pred. No.:      299      Length:      759
Score:          42.00     Matches:      8
Percent Similarity: 90.91%  Conservative: 2
Best Local Similarity: 77.73%  Mismatches:  1
Query Match:     77.78%     Indels:      0
DB:              12       Gaps:        0

```

```

US-09-823-649a-3 (1-11) x BG124823 (1-759)
QY      1 leuserglnluleaialleprotyrglugu 11
|||||.....|
Db      64 TTGTCCACACACACATCCATTCACAGG 32

RESULT 15
AV043104      198 bp  mRNA  linear  EST 23-NOV-1999
LOCUS      AV043104 Mus musculus adult C57BL/6J testis Mus musculus cDNA clone
DEFINITION  1700026C12, mRNA sequence.
ACCESSION  AV043104
VERSION    AV043104.2  GI:4862769
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE  1 (bases 1 to 198)
AUTHORS   Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Alizawa, K.,
            Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
            A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,
            Kikuchi, N., Koike, Y., Matsuyama, T., Mitsuma, H., Oda, H., Owa, C.,
            Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara
            Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomiwa, N.,
            Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
            Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
            RIKEN Mouse ESTs
            Unpublished (1999)
JOURNAL
COMMENT    Contact: Chie Owa
            Genome Science Laboratory
            RIKEN
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-9145
            Fax: 81-298-36-9098
            Email: genome-reset@riken.go.jp
            Thermostabilization and thermocactivation of thermostable enzymes by
            lrehalose and its application for the synthesis of full length cDNA
            (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
            Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
            Please visit our web site (http://genome.riken.go.jp) for
            further details.
            Location/Qualifiers
                source
                1..198
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="1700026C12"
                /clone_lib="Mus musculus adult C57BL/6J testis"
                /sex="male"
                /tissue_type="testis"
                /dev_stage="adult"
                /note="Organ: mammary gland; Vector: p773D-Pac (Pharmacia
                ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                RI; 1st strand cDNA was primed with a Not I - o119g(drr)
                primer [5',
                TGTATCAATCTCACTGGACCGCCGCGAAGCTTTTCTTTTCTTTTCTTTT
                T 3']; double-stranded cDNA was ligated to Eco RI
                adaptors (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of the modified p773 vector.
                RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                constructed and normalized by Bento Soares and M. Patricia
                Bonaldo."
BASE COUNT      61 a      48 c      33 g      56 t
ORIGIN
Alignment Scores:
Pred. No.:      74.3      Length:      198
Score:          41.00     Matches:      8
Percent Similarity: 88.89%  Conservative: 0
Best Local Similarity: 88.89%  Mismatches:  1
Query Match:     75.93%     Indels:      0

```


DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x AV043104 (1-198)

OY 3 GInGIULeuaIaIleProTYrGIuGIu 11

Db 17 CAAGAACTCGCCTATCCTTATGAGAG 43

Search completed: January 21, 2003, 12:24:16
Job time : 1171.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:02:09 ; Search time 28.1429 Seconds

(without alignments)
52.083 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSTPYEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	96.0	11	AA847794	Native DNA polymer
2	48	96.0	11	AA848260	Native DNA polymer
3	48	96.0	11	AA848263	Native DNA polymer
4	48	96.0	11	AA848265	Native DNA polymer
5	48	96.0	545	13 AAR23157	Mutant thermostabl
6	48	96.0	632	13 AAR23156	Mutant thermostabl
7	48	96.0	680	13 AAR23155	Mutant thermostabl
8	48	96.0	758	13 AAR23154	Mutant thermostabl
9	48	96.0	788	13 AAR23153	Mutant thermostabl
10	48	96.0	830	13 AAR23152	Mutant thermostabl

11	48	96.0	830	16 AAR76060	Tfil DNA-polymeras
12	48	96.0	831	16 AAR64273	T. flavus DNA-poly
13	48	96.0	831	19 AAR79961	Thermus flavus DNA
14	48	96.0	831	20 AAR80428	DNA polymerase enz
15	48	96.0	845	16 AAR72875	TspA7 polymerase.
16	45	90.0	11	23 AAR47793	Native DNA polymer
17	45	90.0	11	23 AAM48259	Native DNA polymer
18	45	90.0	11	23 AAM48261	Native DNA polymer
19	45	90.0	11	23 AAM48262	Native DNA polymer
20	45	90.0	11	23 AAM48264	Native DNA polymer
21	45	90.0	528	17 AAR96267	Mutant Thermus aqu
22	45	90.0	544	13 AAR23145	Mutant Thermostabl
23	45	90.0	544	13 AAR23163	Mutant thermostabl
24	45	90.0	544	13 AAR23168	Mutant thermostabl
25	45	90.0	552	22 AAE09304	Thermus aquaticus
26	45	90.0	553	21 AAY44352	Thermoslabile DNA d
27	45	90.0	554	16 AAR66209	Novel thermostable
28	45	90.0	559	21 AAY44353	Thermoslabile DNA d
29	45	90.0	560	17 AAR95544	Thermus flavus DNA
30	45	90.0	560	17 AAR96205	DNA polymerase I e
31	45	90.0	561	18 AAM09315	Mutant Taq polymer
32	45	90.0	561	18 AAM09317	Mutant Taq polymer
33	45	90.0	561	20 AAY00886	Py2 polymerase. T
34	45	90.0	561	20 AAY00888	Py4 polymerase. T
35	45	90.0	562	18 AAM09316	Mutant Taq polymer
36	45	90.0	562	20 AAY00887	Py3 polymerase. T
37	45	90.0	562	22 AAE09305	Thermus thermophil
38	45	90.0	562	22 AAE09310	Thermus aquaticus
39	45	90.0	597	17 AAR99543	Thermus flavus DNA
40	45	90.0	597	17 AAR96204	DNA polymerase I e
41	45	90.0	604	22 AAU00574	Chimeric Cautaq DN
42	45	90.0	604	22 AAU00575	Chimeric Cautaq DN
43	45	90.0	631	13 AAR23144	Mutant thermostabl
44	45	90.0	632	13 AAR23162	Mutant thermostabl
45	45	90.0	632	13 AAR23167	Mutant thermostabl

ALIGNMENTS

RESULT 1	
AA847794	AA847794 standard; peptide: 11 AA.
ID	AA847794
AC	AA847794;
XX	
DT	25-MAR-2002 (first entry)
XX	
DE	Native DNA polymerase motif #4.
XX	
XX	DNA polymerase: reverse transcription; primer: divalent cation; mutant;
KW	transverse transcription reaction; fluorescein; cyanine; thermocactive;
KM	dye; amplification.
XX	
OS	Thermus sp.
XX	
FH	Key
FT	Misc-difference 3 Location/Qualifiers
FT	/label= Gln, Gly
XX	
XX	EP1152062-A2.
XX	
PD	07-NOV-2001.
XX	
PF	12-APR-2001; 2001EP-0109341.
XX	
PR	18-APR-2000; 2000US-198336P.
XX	
PA	(HOFF) HOFFMANN IA ROCHE & CO AG F.
XX	
PI	Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI	Schoenbrunner NJ, Wang AM;
XX	

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-033-297-12
Alignment Scores:
Pred. No.: 0.559 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-2 (1-11) x US-10-033-297-12 (1-1600)
QY 1 Leuser***GluLeu***IleProTyrgluGlu 11
Db 1133 CTCCTCCAGAGCTACGCTACCTTACGAGAG 1165
RESULT 2
US-10-081-806-12
Sequence 12, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-081-806-12
Alignment Scores:
Pred. No.: 0.559 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-2 (1-11) x US-10-081-806-12 (1-1600)
QY 1 Leuser***GluLeu***IleProTyrgluGlu 11
Db 1133 CTCCTCCAGAGCTACGCTACCTTACGAGAG 1165
RESULT 3
US-10-033-297-2
Sequence 2, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Brow, Mary Ann D.
Mast, Andrea L.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-No. US20020187486A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072

FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-033-297-2

Alignment Scores:
Pred. No.: 0.908 Length: 2496
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-2 (1-11) x US-10-033-297-2 (1-2496)

QY 1 Leuser***Gluleu***IleProTyrGlnGlu 11
||||| ||||| ||||| ||||| |||||
Db 2029 CTCCTCGGAGGAGCTTCCATCCCTACGAGAG 2061

RESULT 4
US-10-081-806-2
Sequence 2, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996

APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-081-806-2

Alignment Scores:
Pred. No.: 0.908 Length: 2496
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-2 (1-11) x US-10-081-806-2 (1-2496)

QY 1 Leuser***Gluleu***IleProTyrGlnGlu 11
||||| ||||| ||||| ||||| |||||
Db 2029 CTCCTCGGAGGAGCTTCCATCCCTACGAGAG 2061

RESULT 5
US-10-033-297-7
Sequence 7, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386

FILED DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-033-297-7

Alignment Scores:
Pred. No.: 0.91 Length: 2502
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0

US-09-823-649a-2 (1-11) x US-10-033-297-7 (1-2502)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11
||||| ||||| ||||| ||||| |||||
Db 2035 CTCTCCAGGAGCTTGCATCCCTACAGAGAG 2067

RESULT 6
US-10-033-297-21
Sequence 21, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESSES:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996

APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-033-297-21

Alignment Scores:
Pred. No.: 0.91 Length: 2502
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0

US-09-823-649a-2 (1-11) x US-10-033-297-21 (1-2502)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11
||||| ||||| ||||| ||||| |||||
Db 2035 CTCTCCAGGAGCTTGCATCCCTACAGAGAG 2067

RESULT 7
US-10-081-806-7
Sequence 7, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESSES:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-FEB-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-081-806-7
Alignment Scores:
Pred. No.: 0.91 Length: 2502
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-2 (1-11) x US-10-081-806-7 (1-2502)
QY 1 Leuser***GluLeu***IleProTyrGlu 11
DB 2035 CTCTCCAGAGAGCTTCCATCCCTACGAGAG 2067
RESULT 8
US-10-081-806-21
Sequence 21, Application US/10081806
Publication No.: US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2502 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-081-806-21
Alignment Scores:
Pred. No.: 0.91 Length: 2502
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-2 (1-11) x US-10-081-806-21 (1-2502)
QY 1 Leuser***GluLeu***IleProTyrGlu 11
DB 2035 CTCTCCAGAGAGCTTCCATCCCTACGAGAG 2067
RESULT 9
US-10-033-297-3
Sequence 3, Application US/10033297
Publication No.: US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-MAY-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/582,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2504 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-033-297-3

Alignment Scores:
Pred. No.: 0.911 Length: 2504
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0
DB: 9

US-09-823-649a-2 (1-11) x US-10-033-297-3 (1-2504)

QY 1 Leuser***GluLeu***IleProTyrGlu 11
Db 2038 CTCCTCCAGAGAGCTTCACCTCCCTACGAGAG 2070

RESULT 10
US-10-081-806-3
; Sequence 3, Application US/10081806
; Publication No. US20020197623A1
; GENERAL INFORMATION:
; APPLICANT: Prudent, James R.
; Hall, Jeff G.
; Tyamichev, Victor I.
; TITLE OF INVENTION: Invasive Cleavage of Nucleic Acids
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/081,806
; FILING DATE: 22-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,386
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2504 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-081-806-3

Alignment Scores:
Pred. No.: 0.911 Length: 2504
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0
DB: 9

US-09-823-649a-2 (1-11) x US-10-081-806-3 (1-2504)

QY 1 Leuser***GluLeu***IleProTyrGlu 11
Db 2038 CTCCTCCAGAGAGCTTCACCTCCCTACGAGAG 2070

RESULT 11
US-10-033-297-65
; Sequence 65, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; Tyamichev, Victor I.
; Mast, Andrea L.
; Brow, Mary Ann D.
; TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
; Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/033,297
; FILING DATE: 12-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/350,597
; FILING DATE: 09-Jul-1999
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-Mar-1997
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-Jan-1997
; APPLICATION NUMBER: US 08/759,048
; FILING DATE: 02-Dec-1996
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-Dec-1996
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-Nov-1996
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-Jul-1996
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-033-297-65
Alignment Scores:
Pred. No.: 0.911 Length: 2505
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0
DB: 9
US-09-823-649A-2 (1-11) x US-10-033-297-65 (1-2505)
QY 1 Leuser***Gluleu***IleProTyRGluGln 11
||||| ||||| ||||| ||||| |||||
Db 2035 CTCCTCCAGAGCTAGCCATCCTCTACGAGAG 2067
RESULT 12
US-10-033-297-68
Sequence 68, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES* 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FOS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-033-297-68
Alignment Scores:
Pred. No.: 0.911 Length: 2505
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0
DB: 9
US-09-823-649A-2 (1-11) x US-10-033-297-68 (1-2505)
QY 1 Leuser***Gluleu***IleProTyRGluGln 11
||||| ||||| ||||| ||||| |||||
Db 2035 CTCCTCCAGAGCTAGCCATCCTCTACGAGAG 2067
RESULT 13
US-10-033-297-70
Sequence 70, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2499
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-033-297-70
Alignment Scores:
Pred. No.: 0.911 Length: 2505
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-2 (1-11) x US-10-033-297-70 (1-2505)
QY 1 LeuSer***Glueu***lleProTyrglu 11
||||| ||||| - ||||| |||||
Db 2032 CTCCTCCGAGAGCTAGCATCCCTTACGAGAG 2067
RESULT 14
US-10-033-297-1
Sequence 1, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NO. US20020187486A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314

FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-033-297-1
Alignment Scores:
Pred. No.: 0.912 Length: 2506
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 9 Gaps: 0
US-09-823-649a-2 (1-11) x US-10-033-297-1 (1-2506)
QY 1 LeuSer***Glueu***lleProTyrglu 11
||||| ||||| ||||| |||||
Db 2032 CTCCTCCGAGAGCTAGCATCCCTTACGAGAG 2064
RESULT 15
US-10-081-806-1
Sequence 1, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.

```

?      REGISTRATION NUMBER: 40,027
?      REFERENCE /DOCKET NUMBER: FOR-02564
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (415) 705-8410
?      TELEFAX: (415) 397-8338
?      INFORMATION FOR SEQ ID NO: 1 :
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 2506 base pairs
?          TYPE: nucleic acid
?          STRANDEDNESS: double
?          TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?      SEQUENCE DESCRIPTION: SEQ ID NO: 1 :
US-10-081-806-1

Alignment Scores:
Pred. No.:           0.912             Length:       2506
Score:              44.00             Matches:      9
Percent Similarity: 81.82%            Conservative: 0
Best Local Similarity: 81.82%         Mismatches:   2
Query Match:        93.62%            Indels:       0
                9                      Gaps:         0

US-09-823-649A-2 (1-11) x US-10-081-806-1 (1-2506)
OY      1 leuSer****GluIeu****lleProTyrCjUglu 11
        ||||| | ||||| ||||| ||||| |||||
Db      2032 CTTCTCCAGGAGACTACGCATTCCTTAGCAGAG 2064

```

Search completed: January 21, 2003, 10:07:41
Job time : 31.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649a-2
Perfect score: 47
Sequence: 1 LSXELXIPYEE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -s -
-O/cogn2.1/USPRO.spool/US09823649/unal.21012003.093151.24616/app.query.fasta.1.1393
-MODE=frame+ p2n.model -DEV=xlp
-UNIT=bits -START=1 -END=1 -MATRIX=D10sum62 -TRANS=human40.cdi -LIST=45
-DB=EST_QFMT-fastcap -SUFFIX=rst -MINMATCH=0.1 -DOOPCL=0 -DOOPEXT=0
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.ecgn.1.1.2820.gunal.21012003.093151.24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6 -YGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST.*
1: em_estba.*
2: em_estlum.*
3: em_estlin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vit.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	87.2	447	13	B1929077
2	41	87.2	480	12	BF050758
3	41	87.2	495	13	H1129400
4	41	87.2	507	9	A1895448
5	41	87.2	534	14	B0489763
6	41	87.2	592	10	BE434433
7	41	87.2	629	14	B0990012
8	41	87.2	728	14	B0856219
9	41	87.2	768	13	BM411779
10	40	85.1	658	14	B0404238
11	38	80.9	183	10	AV415207
12	38	80.9	249	9	AV035497
13	38	80.9	411	10	AV428667
14	38	80.9	416	12	BF758509
15	38	80.9	426	10	AV415722
16	37	78.7	231	9	AV036702
17	37	78.7	270	9	AV037125
18	37	78.7	274	9	AV036809
19	37	78.7	279	9	AV035994
20	37	78.7	281	9	AV037570
21	37	78.7	284	9	AV035331
22	37	78.7	298	9	AV034428
23	37	78.7	306	9	AV225900
24	37	78.7	319	9	AV036154
25	37	78.7	345	9	AT152562
26	37	78.7	516	13	BM249366
27	37	78.7	562	13	BM250725
28	37	78.7	682	14	B0185049
29	37	78.7	788	12	BG216020
30	36	76.6	274	9	AV223368
31	36	76.6	438	13	BM482565
32	36	76.6	571	12	BG844832
33	36	76.6	587	13	RJ315731
34	36	76.6	721	13	BJ321189
35	36	76.6	760	13	BM291026
36	35	74.5	274	9	AV223298
37	35	74.5	292	9	AV034586
38	35	74.5	353	13	B1680984
39	35	74.5	440	9	A1850988
40	35	74.5	482	9	AA701609
41	35	74.5	493	14	R49433
42	35	74.5	565	10	BE656341
43	35	74.5	647	12	BF701996
44	35	74.5	650	17	AZ826009
45	35	74.5	664	14	BQ139058

ALIGNMENTS

Result	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
1	B1929077	B1929077	B1929077	B1929077	B1929077.1	GI:16241891	tomato.
1	B1929077	EST548966 tomato flower, 3 - 8 mm buds	447 bp	mrna	linear	EST 18-OCT-2001	Lycopersicon esculentum
1	B1929077	clone cTOB27117 5' end, mRNA sequence.					Lycopersicon esculentum cDNA
1	B1929077	EST.					Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
1	B1929077	1 (bases 1 to 447)					

AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Karamychera, S.A., Tsai, J.,
 Uiterback, T., Van Aken, S., Roming, C.M., Nierman, W., Fraser, C.M.,
 Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute

FEATURES Seq primer: T3.

source Location/Qualifiers

1..447
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="G0827117"
 /clone_lib="tomato flower, 3 - 8 mm buds"
 /tissue_type="flower"
 /dev_stage="3-8mm buds"
 /note="Vector: pBluescript SK(-); site: 1: EcoRI; site: 2:
 XhoI; supplier: Cornell University; sequencing: The
 Institute for Genomic Research. Flower buds and flowers
 were taken from greenhouse plants (4-8 wks old, TA496).
 They were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."
BASE COUNT 118 a 92 c 89 g 148 t
ORIGIN

Alignment Scores:

Pred. No.: 96.4 % **Length:** 447
Score: 41.00 **Matches:** 7
Percent Similarity: 81.82% **Conservative:** 2
Best Local Similarity: 63.64% **Mismatches:** 2
Query Match: 87.23% **Indels:** 0
DB: 13 **Gaps:** 0

US-09-823-649a-2 (1-11) x B1929077 (1-447)

QY 1 Leuser***Gluuu***IleProTYrGluGlu 11
 ||||| ||||| ::||| ||||| ::|||
Db 331 CTCGCGCTGAGCTCTCAATTCGTAAGAGAT 363

RESULT 2
LOCUS BF050758 480 bp mRNA linear EST 18-MAY-2001
DEFINITION ST435916 tomato developing/immature green fruit Lycopersicon
 esculentum cDNA clone cLEM19022 5' sequence, mRNA sequence.
ACCESSION BF050758
VERSION BF050758.1 GI:10804654
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE 1 (bases 1 to 480)
 Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Roming, C.M.,
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley,
 S.D.
AUTHORS

TITLE Generation of ESTs from tomato fruit tissue, immature green
 Unpublished (2000)
JOURNAL Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
COMMENT Email: <http://www.genome.clemson.edu/orders/index.html>.
 Location/Qualifiers

FEATURES 1..480
source

/organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLEM19022"
 /clone_lib="tomato developing/immature green fruit"
 /tissue_type="fruit"
 /dev_stage="immature green (5-35 days post-anthesis)"
 /note="Vector: pBluescriptSKmcdap1; site: 1: EcoRI;
 site: 2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
 harvested at 7 day intervals through 35 dpa. Equal masses
 of tissue from each stage were combined (including seeds
 and locules) prior to mRNA isolation."
BASE COUNT 116 a 102 c 119 g 143 t
ORIGIN

Alignment Scores:

Pred. No.: 106 **Length:** 480
Score: 41.00 **Matches:** 7
Percent Similarity: 81.82% **Conservative:** 2
Best Local Similarity: 63.64% **Mismatches:** 2
Query Match: 87.23% **Indels:** 0
DB: 12 **Gaps:** 0

US-09-823-649a-2 (1-11) x BF050758 (1-480)

QY 1 Leuser***Gluuu***IleProTYrGluGlu 11
 ||||| ||||| ::||| ||||| ::|||
Db 261 CTCGCGCTGAGCTCTCAATTCGTAAGAGAT 293

RESULT 3
LOCUS B1129400 495 bp mRNA linear EST 31-DEC-2001
DEFINITION G090P24Y Populus cambium cDNA library Populus tremula x Populus
 tremuloides cDNA, mRNA sequence.
ACCESSION B1129400
VERSION B1129400.1 GI:18013371
KEYWORDS EST.
SOURCE Populus tremula x Populus tremuloides.
ORGANISM Populus tremula x Populus tremuloides
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
 1 (bases 1 to 495)
 Hertzberg, M., Aspeborg, H., Erlandsson, R., Bjorkbacka, H., Hilltonen
 T., Karlsson, J., Teeri, T., Gustafsson, P., Bahler, R., Jansson, S.,
 Nilsson, O., Sundberg, B., Nilsson, P., Uhlen, M., Sandberg, G. and
 Lundberg, T.
AUTHORS Gene expression in Populus
 Unpublished (2001)
TITLE Contact: Erlandsson R
 Department of Biotechnology
 Royal Institute of Technology
 Teknikringen 30, Stockholm S-10044, Sweden
 Tel: 46 8 790 8287
 Fax: 46 8 245452
 Email: rikeri@biochem.kth.se.
COMMENT Location/Qualifiers

FEATURES 1..495
 /organism="Populus tremula x Populus tremuloides"
 /db_xref="taxon:47664"
 /clone_lib="Populus cambium cDNA library"
 /note="Organ: cambium"

BASE COUNT 109 a 146 c 117 g 123 t
ORIGIN

Alignment Scores:

Pred. No.: 110 **Length:** 495
Score: 41.00 **Matches:** 7
Percent Similarity: 81.82% **Conservative:** 2
Best Local Similarity: 63.64% **Mismatches:** 2
Query Match: 87.23% **Indels:** 0
DB: 13 **Gaps:** 0

US-09-823-649A-2 (1-11) x B1129400 (1-495)

QY 1 leuser***Gluleu***lleprotyrgluciu 11
||||| ||||| ::||| ||||| ::|||

Db 59 CTATCTGGCGAGTGTCCGTTCCGTATGAGAT 91

RESULT 4

AI895448

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

tomato.

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 507)

Alcala,J., Vredalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,

Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ann,S., Rønning,

C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence"

Location/Qualifiers

1..507

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CUEC7H2"

/clone_lib="tomato callus, TAMU"

/tissue_type="callus"

/dev_stage="25-40 days old"

/lab_host="XLI-Blue MRF"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; supplier: Giovannoni laboratory; CUGC - Cotyledons

at both ends and placed on MS medium with no selection. cut

Mixed callus was harvested at 25 and 40 days and included

undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 125 a 138 c 114 g 130 t

ORIGIN

Alignment Scores:

Pred. No.: 114

Score: 41.00

Percent Similarity: 81.82%

Best Local Similarity: 63.64%

Query Match: 87.23%

DB: 9

Gaps: 0

US-09-823-649A-2 (1-11) x AI895448 (1-507)

QY 1 leuser***Gluleu***lleprotyrgluciu 11

||||| ||||| ::||| ||||| ::|||

Db 66 CMTCCGCTGAGCTCTCAGTTCGATGACGAT 98

RESULT 5

BO489763

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

tomato.

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 592)

Alcala,J., Vredalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,

Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Rønning,C.M.,

Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley

S.D. Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished (2000)

Location/Qualifiers

1..592

/organism="Beta vulgaris"

/db_xref="taxon:161934"

/clone="1-1-9"

/clone_lib="Sugar beet MP12-ADIS-006 Lambda zap II

library"

/dev_stage="4 week old pot-grown plants"

/note="Organ: shoot and root; Vector: pBluescript SK- from

lambda zap II; cDNA (lambda zap-II) library from sugar

beet, whole plant mRNA, prepared using the Stratagene

UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites

and orientation:

rev="3'-SacI-SK-ECORI-CGCACGACG-5'p-cDNA-polya-XhoI-KpnI-17

unit"

BASE COUNT 143 a 112 c 127 g 152 t

ORIGIN

Alignment Scores:

Pred. No.: 122

Score: 41.00

Percent Similarity: 81.82%

Best Local Similarity: 63.64%

Query Match: 87.23%

DB: 14

Gaps: 0

US-09-823-649A-2 (1-11) x BO489763 (1-592)

QY 1 leuser***Gluleu***lleprotyrgluciu 11

||||| ||||| ::||| ||||| ::|||

Db 281 CTCTCTGAGTCTCTCAGTTCGATGACGAT 43

RESULT 6

BE434433

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

tomato.

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 592)

Alcala,J., Vredalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,

Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Rønning,C.M.,

Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley

S.D. Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished (2000)

Location/Qualifiers

1..592

/organism="Beta vulgaris"

/db_xref="taxon:161934"

/clone="1-1-9"

/clone_lib="Sugar beet MP12-ADIS-006 Lambda zap II

library"

/dev_stage="4 week old pot-grown plants"

/note="Organ: shoot and root; Vector: pBluescript SK- from

lambda zap II; cDNA (lambda zap-II) library from sugar

beet, whole plant mRNA, prepared using the Stratagene

UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites

and orientation:

rev="3'-SacI-SK-ECORI-CGCACGACG-5'p-cDNA-polya-XhoI-KpnI-17

unit"

BASE COUNT 143 a 112 c 127 g 152 t

ORIGIN

Alignment Scores:

Pred. No.: 122

Score: 41.00

Percent Similarity: 81.82%

Best Local Similarity: 63.64%

Query Match: 87.23%

DB: 14

Gaps: 0

US-09-823-649A-2 (1-11) x BO489763 (1-592)

QY 1 leuser***Gluleu***lleprotyrgluciu 11

||||| ||||| ::||| ||||| ::|||

Db 281 CTCTCTGAGTCTCTCAGTTCGATGACGAT 43

RESULT 6

BE434433

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

tomato.

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 592)

Alcala,J., Vredalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,

Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Rønning,C.M.,

Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley

S.D. Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished (2000)

Location/Qualifiers

1..592

/organism="Beta vulgaris"

/db_xref="taxon:161934"

/clone="1-1-9"

/clone_lib="Sugar beet MP12-ADIS-006 Lambda zap II

library"

/dev_stage="4 week old pot-grown plants"

/note="Organ: shoot and root; Vector: pBluescript SK- from

lambda zap II; cDNA (lambda zap-II) library from sugar

beet, whole plant mRNA, prepared using the Stratagene

UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites

and orientation:

rev="3'-SacI-SK-ECORI-CGCACGACG-5'p-cDNA-polya-XhoI-KpnI-17

unit"

BASE COUNT 143 a 112 c 127 g 152 t

ORIGIN

Alignment Scores:

Pred. No.: 122

Score: 41.00

Percent Similarity: 81.82%

Best Local Similarity: 63.64%

Query Match: 87.23%

DB: 14

Gaps: 0

US-09-823-649A-2 (1-11) x BO489763 (1-592)

QY 1 leuser***Gluleu***lleprotyrgluciu 11

||||| ||||| ::||| ||||| ::|||

Db 281 CTCTCTGAGTCTCTCAGTTCGATGACGAT 43

RESULT 6

BE434433

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

tomato.

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 592)

Alcala,J., Vredalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,

Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Rønning,C.M.,

Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley

S.D. Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished (2000)

Location/Qualifiers

1..592

/organism="Beta vulgaris"

/db_xref="taxon:161934"

/clone="1-1-9"

/clone_lib="Sugar beet MP12-ADIS-006 Lambda zap II

library"

/dev_stage="4 week old pot-grown plants"

/note="Organ: shoot and root; Vector: pBluescript SK- from

lambda zap II; cDNA (lambda zap-II) library from sugar

beet, whole plant mRNA, prepared using the Stratagene

UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites

and orientation:

rev="3'-SacI-SK-ECORI-CGCACGACG-5'p-cDNA-polya-XhoI-KpnI-17

unit"

BASE COUNT 143 a 112 c 127 g 152 t

ORIGIN

Alignment Scores:

Pred. No.: 122

Score: 41.00

Percent Similarity: 81.82%

Best Local Similarity: 63.64%

Query Match: 87.23%

DB: 14

Gaps: 0

US-09-823-649A-2 (1-11) x BO489763 (1-592)

QY 1 leuser***Gluleu***lleprotyrgluciu 11

||||| ||||| ::||| ||||| ::|||

Db 281 CTCTCTGAGTCTCTCAGTTCGATGACGAT 43

RESULT 6

BE434433

COMMENT

Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

source
Location/Qualifiers

1..592
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEB17C9"

/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"

/note="Vector: pBluescriptSKm(+/-) Site 1: EcoRI;
Site 2: XhoI; fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

BASE COUNT 140 a 149 c 137 g 166 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
139	592	7	2	2	0	0
Percent Similarity:	41.00	81.82%	63.64%	87.23%		
Best Local Similarity:	63.64%					
Query Match:	87.23%					

US-09-823-649a-2 (1-11) x BE954433 (1-592)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 194 CTCCTCGCTGAGCTCTCAGTTCCGATATGAGCAT 226

RESULT 7

LOCUS

BQ990012/c 629 bp mRNA linear EST 21-AUG-2002

DEFINITION

OGF19P10.yg.ab1 OG_EFGH1 lettuce serriola lactuca sativa cDNA clone

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGF19P10"

/clone_lib="OG_EFGH1 lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBluescriptSKm(+/-) The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.clemson.edu/TAG_TISSUE-flowers-post-fertilized

TAG_TISSUE-flowers-post-fertilized

TAG_SEQ-TAGCATCGAGC"

BASE COUNT 138 a 137 c 183 g 170 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

Percent Similarity:

Best Local Similarity:

Query Match:

US-09-823-649a-2 (1-11) x BQ990012 (1-629)

QY 1 Leuser***GluLeu***IleProtyrGluGlu 11

DB 471 CTCCTGAGCAGACTTCCTCCGATGAGCAT 439

RESULT 8

LOCUS

```

/lab_host="E.coli"
/note="vector: pBKCMASf1AB: The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cspdb.ucdavis.edu/
TAG_LIB-06_ABCDI lettuce salinas
TAG_TISSUE="chemical" induction
TAG_SEQ="GTAACCGCGC"

BASE COUNT      203 a      203 c      153 g      169 t

ORIGIN

Alignment Scores:
Pred. No.:      182      length:      728
Score:          41.00      Matches:      7
Percent Similarity: 81.82%      Conservative: 2
Best Local Similarity: 63.64%      Mismatches: 2
Query Match:    87.23%      Indels:      0
DB:             14      Gaps:        0

US-09-823-649A-2 (1-11) x BQ856219 (1-728)

OY      1 LeuSer***GlutLeu***IleProTyrGluGlu 11
|||||  |||||  ::|||  |||||  ::|||
Db      299 CTTTCAGCAGACTTCGCTTCGATGAGAT 331

RESULT 9
BMA11779      -      *      768 bp      mRNA      linear      EST 22-JAN-2002
LOCUS      EST586106 tomato breaker fruit Lycopersicon esculentum cDNA clone
DEFINITION      CLEBSV16 5' end, mRNA sequence.
ACCESSION      BMA11779
VERSION      BMA11779.1 GI:18263409
KEYWORDS      EST.
SOURCE      tomato.
ORGANISM      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 768)
Alcala,J., Vrebalov,J., White,R., Vision,T., Karanaycheva,S.A., Tsai
,J., Bougri,O., Kirkness,E., Ullrichback,T., Van Aken,S., Roming
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..768
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEBSV16"
/clone_lib="tomato breaker fruit"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKM(udapt: Site1: EcoRI;
Site2: XhoI; supplier: Boyce Thompson Institute;
Sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit

```

```

were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

BASE COUNT      209 a      181 c      173 g      205 t

ORIGIN

Alignment Scores:
Pred. No.:      195      length:      768
Score:          41.00      Matches:      7
Percent Similarity: 81.82%      Conservative: 2
Best Local Similarity: 63.64%      Mismatches: 2
Query Match:    87.23%      Indels:      0
DB:             13      Gaps:        0

US-09-823-649A-2 (1-11) x BMA11779 (1-768)

OY      1 LeuSer***GlutLeu***IleProTyrGluGlu 11
|||||  |||||  ::|||  |||||  ::|||
Db      261 CTTTCAGCAGACTTCGCTTCGATGAGAT 293

RESULT 10
BQ404238      -      *      658 bp      mRNA      linear      EST 22-MAY-2002
LOCUS      GA_Ed0068A08f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION      GA_Ed0068A08f Gossypium arboreum 7-10 dpa fiber library Gossypium
arborescens cDNA clone GA_Ed0068A08f, mRNA sequence.
ACCESSION      BQ404238
VERSION      BQ404238.1 GI:21091925
KEYWORDS      EST.
SOURCE      Gossypium arboreum.
ORGANISM      Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 658)
Wing,R.A., Preisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total High Quality bases - 607
Seq primer: TAATCACTCACTATACGG
High quality sequence stop: 656.
Location/Qualifiers
1..658
/organism="Gossypium arboreum"
/cultivar="FARA"
/db_xref="taxon:29729"
/clone="GA_Ed0068A08f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBK-CMV; Site1: EcoRI; Site2: XhoI"

BASE COUNT      170 a      164 c      122 g      201 t

ORIGIN

Alignment Scores:
Pred. No.:      256      length:      658
Score:          40.00      Matches:      7
Percent Similarity: 81.82%      Conservative: 2
Best Local Similarity: 63.64%      Mismatches: 2
Query Match:    85.11%      Indels:      0
DB:             14      Gaps:        0

US-09-823-649A-2 (1-11) x BQ404238 (1-658)

```


TITLE Generation of 7137 non-redundant expressed sequence tags from a legume, *Lotus japonicus*

JOURNAL DNA Res. 7 (2), 127-130 (2000)

MEDLINE 20277479

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/

FEATURES
source
1..411
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone_lib="MM10503_r"
/dev_stage="Young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; isolate=Miyakojima MG-20"
2 others

BASE COUNT 86 a 143 c 90 g 90 t

ORIGIN

Alignment Scores:
Pred. No.: 358 Length: 411
Score: 38.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 80.85% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-2 (1-11) x AV428667 (1-411)

QY 1 LeuSer**GluLeu**IleProTyGluGlu 11
||||| ||||| - *::||| ||||| |||||

Db 374 CTCTCGGAGGAGCTGCTGTCGATGATGAA 406

RESULT 14
BF758509/c 416 bp mRNA linear EST 12-JAN-2001

LOCUS BF758509

DEFINITION MR4-CT0538-141100-104-b08 CT0538 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF758509

VERSION BF758509.1 GI:12106409

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 416)
Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Britones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, R.F., Goldman, G.H., Carvalhal, A.F., Matsukuma, A., Hata, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jonsson, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&t2=MR4-CT0538-141100-104-b08&t3=2000-11-14&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 88.
Location/Qualifiers

source
1..416
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0538"
/dev_stage="Adult"

BASE COUNT 129 a 106 c 85 g 96 t

ORIGIN

Alignment Scores:
Pred. No.: 363 Length: 416
Score: 38.00 Matches: 8
Percent Similarity: 72.73% Conservative: 0
Best Local Similarity: 72.73% Mismatches: 3
Query Match: 80.85% Indels: 0
DB: 12 Gaps: 0

US-09-823-649a-2 (1-11) x BF758509 (1-416)

QY 1 LeuSer**GluLeu**IleProTyGluGlu 11
||||| ||||| ||||| ||||| |||||

Db 182 TTGCTGTGAGGAGGAGCTATCCCTATGAGAG 150

RESULT 15
AV415722 426 bp mRNA linear EST 23-MAY-2000

LOCUS AV415722

DEFINITION AV415722 Lotus japonicus young plants (two-week old) Lotus japonicus cDNA clone MM115e12_r 5', mRNA sequence.

ACCESSION AV415722

VERSION AV415722.1 GI:7744898

KEYWORDS EST.

SOURCE Lotus japonicus.

ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Lotiae;
Lotus.

REFERENCE 1 (bases 1 to 426)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Generation of 7137 non-redundant expressed sequence tags from a legume, *Lotus japonicus*

TITLE Generation of 7137 non-redundant expressed sequence tags from a legume, *Lotus japonicus*

JOURNAL DNA Res. 7 (2), 127-130 (2000)

MEDLINE 20277479

COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/

FEATURES
source
1..426
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone_lib="MM115e12_r"
/dev_stage="Young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; isolate=Miyakojima MG-20"
94 t

BASE COUNT 90 a 147 c 95 g 94 t

ORIGIN

Alignment Scores:
Pred. No.: 375 Length: 426
Score: 38.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 80.85% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-2 (1-11) x AV415722 (1-426)

QY 1 LeuSer***GluLeu***1leProTyrGlu 11
||||| ||||| ::|||
Db 375 CTCCTCGGCGAGCTCTGTGTTCGTATGATGAA 407

Search completed: January 21, 2003, 12:24:13
Job time : 1171.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5114 Seconds
(without alignments)
165,854 Million cell updates/sec

Title: US-09-823-649a-6
Perfect score: 49
Sequence: 1 LSKRIGLSYSE 11

Scoring table:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 393868 segs, 222934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODE=frame_p2n_model -DEV=xlp
-Q/cgnt2.1/usfpo_spool/US09823649/runat_21012003.093151.24626/app_query.fasta_1.1393
-DB=publshed_applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62
-TRN=human40 cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALJN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649.@CGN1.1.77.@runat_21012003.093151.24626
-NCPU=6 -ICPU=3 -NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV=TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA.*

1: /cgnt2.6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgnt2.6/prodata/1/pubpna/US06_PUB_PUB.seq.*
3: /cgnt2.6/prodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgnt2.6/prodata/1/pubpna/US07_PUB_PUB.seq.*
5: /cgnt2.6/prodata/1/pubpna/US07_PUBCOMB.seq.*
6: /cgnt2.6/prodata/1/pubpna/US08_PUB_PUB.seq.*
7: /cgnt2.6/prodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgnt2.6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgnt2.6/prodata/1/pubpna/US09_PUB_PUB.seq.*
10: /cgnt2.6/prodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgnt2.6/prodata/1/pubpna/US10_PUB_PUB.seq.*
12: /cgnt2.6/prodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgnt2.6/prodata/1/pubpna/US60_PUB_PUB.seq.*
14: /cgnt2.6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	39	79.6	471 10	US-09-864-761-62
2	39	79.6	668 10	US-09-864-761-16902
3	37	75.5	1314 10	US-09-974-300-1325
4	36	73.5	540 9	US-09-796-692-7302

Result No.	Score	Query Match Length	ID	Description
5	36	73.5	1024 10	US-09-925-301-15
6	35	71.4	193 10	US-09-864-761-26194
7	35	71.4	300 10	US-09-294-0938-3331
8	35	71.4	568 10	US-09-864-761-9842
9	35	71.4	10885 10	US-09-764-877-3843
10	35	71.4	12651 10	US-09-804-4744-3
11	35	71.4	155074 9	US-10-026-188-6
12	35	71.4	397658 10	US-09-813-920-3
13	34	69.4	431 10	US-09-954-456-1202
14	34	69.4	431 10	US-09-880-107-709
15	34	69.4	462 10	US-09-864-761-5243
16	34	69.4	471 10	US-09-604-2878-222
17	34	69.4	471 10	US-09-339-338-222
18	34	69.4	471 12	US-10-007-805-222
19	34	69.4	614 10	US-09-925-300-653
20	34	69.4	1083 10	US-09-974-300-2681
21	34	69.4	1491 10	US-09-808-483-13
22	34	69.4	1833 10	US-09-815-242-6921
23	34	69.4	2176 8	US-08-808-031A-1
24	34	69.4	2729 10	US-09-822-849A-529
25	34	69.4	2751 9	US-09-938-842A-1105
26	34	69.4	25377 9	US-10-061-119-4
27	34	69.4	52354 10	US-09-742-311-3
28	33	67.3	167 10	US-09-864-761-17003
29	33	67.3	219 10	US-09-560-863-409
30	33	67.3	360 10	US-09-910-943-330
31	33	67.3	477 10	US-09-864-761-178
32	33	67.3	480 10	US-09-815-242-3961
33	33	67.3	480 10	US-09-764-887-121
34	33	67.3	507 9	US-10-046-935-494
35	33	67.3	507 9	US-09-878-178-494
36	33	67.3	587 10	US-09-864-761-7966
37	33	67.3	734 10	US-09-910-943-330
38	33	67.3	774 10	US-09-815-242-6584
39	33	67.3	1081 10	US-09-070-927A-773
40	33	67.3	1347 9	US-09-938-842A-268
41	33	67.3	1970 10	US-09-864-761-4957
42	33	67.3	2000 9	US-09-938-842A-3057
43	33	67.3	2769 9	US-09-911-176B-1
44	33	67.3	3072 10	US-09-981-900B-8
45	33	67.3	3209 10	US-09-220-091-4
			3723 12	US-10-001-215-6

ALIGNMENTS

RESULT 1
US-09-864-761-62/c
Sequence 62, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 62
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005301.16
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
US-09-864-761-62
Alignment Scores:
Pred. No.: 4.39 Length: 471
Score: 39.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 79.59% Indels: 0
DB: 10 Gaps: 0
US-09-823-649a-6 (1-11) x US-09-864-761-62 (1-471)
Qy 1 LeusertLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||
Db 241 TTACAAAGCGGTAGGCTCTCAGTTTCAGAG 209
RESULT 2
US-09-864-761-16902/c
; Sequence 16902, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16902
; LENGTH: 668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005301.16
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: SWISSPROT HIT: P51811, EVALU 5.00e-43
; OTHER INFORMATION: NT HIT: q11418230, EVALU 1.00e-100
; OTHER INFORMATION: EST_HUMAN HIT: BE791300.1, EVALU 2.50e+00
US-09-864-761-16902
Alignment Scores:
Pred. No.: 6.66 Length: 668
Score: 39.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 79.59% Indels: 0
DB: 10 Gaps: 0
US-09-823-649a-6 (1-11) x US-09-864-761-16902 (1-668)
Qy 1 LeusertLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||
Db 207 TTACAAAGCGGTAGGCTCTCAGTTTCAGAG 175
RESULT 3
US-09-974-300-1325
; Sequence 1325, Application US/09974300
```

```
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1325
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1314)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-1325

Alignment Scores:
Pred. No.: 41.2 Length: 1314
Score: 37.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 75.51% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-6 (1-11) x US-09-974-300-1325 (1-1314)
QY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:
Db 577 CTGTGAGAGAGAGCGGCGCTTCGCTTCAGAG 609

RESULT 4
US-09-796-692-7302
; Sequence 7302, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
```

```
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7302
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (502)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7302

Alignment Scores:
Pred. No.: 23.8 Length: 540
Score: 36.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.47% Indels: 0
DB: 9 Gaps: 0

US-09-823-649A-6 (1-11) x US-09-796-692-7302 (1-540)
QY 4 ArgIleSylLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:
Db 360 ACGATTGGCCCTGTCAGTTTCAGAG 383

RESULT 5
US-09-925-301-15
; Sequence 15, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antipodides
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (938)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (1005)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (1012)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (1019)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-15

Alignment Scores:
Pred. No.: 50.9 Length: 1024
Score: 36.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.47% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-6 (1-11) x US-09-925-301-15 (1-1024)
```

QY 4 Arg11eclYleuSerValSerClu 11
|||||
Db 444 AGCATGGCTGCTCATGTTTCAGAG 467

RESULT 6
US-09-864-761-26194
; Sequence 26194, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26194
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022312.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELIX, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: NT HIT: AJ277662.1, EVALUATE 1.10e-01
; OTHER INFORMATION: SWISSPROT HIT: P50551, EVALUATE 9.90e-01

; OTHER INFORMATION: EST_HUMAN HIT: BE393892.1, EVALUATE 1.00e-104
US-09-864-761-26194

Alignment Scores:
Pred. No.: 11 6
Score: 35.00
Percent Similarity: 90.91%
Best Local Similarity: 72.73%
Query Match: 71.43%
DB: 10
Gaps: 0

US-09-823-649a-6 (1-11) x US-09-864-761-26194 (1-193)

QY 1 leuSerlysaArg11eclYleuSerValSerClu 11
|||||
Db 89 CTCAGTCGTCGATCGCATGCTCATGTCAGAG 121

RESULT 7
US-09-294-093B-3331/c
; Sequence 3331, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laloudi, Raghu Nath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN YASSIN.
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 3331
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incode ID No. US20010051335A1 700379842H1
; NAME/KEY: unsure
; LOCATION: 93
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3331

Alignment Scores:
Pred. No.: 19 6
Score: 35.00
Percent Similarity: 100.00%
Best Local Similarity: 70.00%
Query Match: 71.43%
DB: 10
Gaps: 0

US-09-823-649a-6 (1-11) x US-09-294-093B-3331 (1-300)

QY 1 leuSerlysaArg11eclYleuSerValSer 10
|||||
Db 210 CTCACAGACGATGGCGCTCCGCTCTCT 181

RESULT 8
US-09-864-761-9842
; Sequence 9842, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

```

? PRIOR APPLICATION NUMBER: US 60/180,312
? PRIOR FILING DATE: 2000-02-04
? PRIOR APPLICATION NUMBER: US 60/207,456
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: US 09/632,366
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: GB 24263,6
? PRIOR FILING DATE: 2000-10-04
? PRIOR APPLICATION NUMBER: US 60/236,359
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: PCT/US01/00666
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 9842
? LENGTH: 568
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AL022312.7
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.4
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 2.3
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.4
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.5
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 2.6
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1.3
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 1.4
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 2.4
US-09-864-761-9842

Alignment Scores:
Pred. No.: 41.9 Length: 568
Score: 35.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-6 (1-11) x US-09-864-761-9842 (1-568)
QY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||:
Db 290 CTCACGCTCGCATCCAGCTCTCTATTGTCACAG 322

RESULT 9
US-09-764-877-3843/c
; Sequence 3843, Application US/09764877
; Patent No. US20020147740A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Yifeng

```

```

? APPLICANT: Rosen et al.
? TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
? FILE REFERENCE: PC005
? CURRENT APPLICATION NUMBER: US/09/764,877
? CURRENT FILING DATE: 2001-01-17
? PRIOR application data removed - refer to PALM or file wrapper
? NUMBER OF SEQ ID NOS: 4031
? SOFTWARE: Patcraft ver. 2.0
? SEQ ID NO 3843
? LENGTH: 10885
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-764-877-3843

Alignment Scores:
Pred. No.: 1.41e+03 Length: 10885
Score: 35.00 Matches: 7
Percent Similarity: 90.00% Conservative: 2
Best Local Similarity: 70.00% Mismatches: 1
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-6 (1-11) x US-09-764-877-3843 (1-10885)
QY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 10
|||||:|||||:|||||:|||||:|||||:
Db 1141 CTCACATAGATGAGATTGGGCTCTCTATTGCT 1112

RESULT 10
US-09-804-474A-3
; Sequence 3, Application US/09804474A
; Patent No. US20020119518A1
; GENERAL INFORMATION:
; APPLICANT: KODET, Suetan et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS.
; FILE REFERENCE: C1000891
; CURRENT APPLICATION NUMBER: US/09/804,474A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: pastseq for windows version 4.0
; SEQ ID NO 3
; LENGTH: 126512
; TYPE: DNA
; ORGANISM: Human
; EVALUATE:
; NAME/KEY: s100c10
; LOCATION: (1)-(126512)
; OTHER INFORMATION: n = A,T,C or G
US-09-804-474A-3

Alignment Scores:
Pred. No.: 2.57e+04 Length: 126512
Score: 35.00 Matches: 7
Percent Similarity: 90.00% Conservative: 2
Best Local Similarity: 70.00% Mismatches: 1
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-6 (1-11) x US-09-804-474A-3 (1-126512)
QY 2 SerLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||:
Db 123466 AATTAAGATGAGTGGCTAGATGCTCTGAC 123495

RESULT 11
US-10-026-186-6
; Sequence 6, Application US/10026188
; Patent No. US20020164645A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Zhang, Yifeng

```



```
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
FILE OF INVENTION: Ion Channel
FILE REFERENCE: 02307E-114910US
CURRENT APPLICATION NUMBER: US/10/026,188
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US 60/259,379
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 155074
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human genomic region containing 11p15.5
OTHER INFORMATION: (Human Chromosome 11p15.5 PAC clone PDU915f1)
OTHER INFORMATION: containing Kv10T1 gene)
US-10-026-188-6
```

```
Alignment Scores:
Pred. No.: 3.26e+04 length: 155074
Score: 35.00 Matches: 7
Percent Similarity: 72.73% Conservative: 1
Best Local Similarity: 63.64% Mismatches: 3
Query Match: 71.43% Indels: 0
DB: 9 Gaps: 0
```

US-09-823-649a-6 (1-11) x US-10-026-188-6 (1-155074)

```
OY 1 LeuSerlysaRg1leGlyleuSerValSerGlu 11
||| |||||:||||| ||| |||
Db 28786 CTAGGAAAGATGGCGCTTGGGGTTGGGGAG 28818
```

```
RESULT 12
US-09-813-320-3
Sequence 3, Application US/09813320
Patent No. US20020142378A1
GENERAL INFORMATION:
APPLICANT: ZHANG, Hongyu et al.
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NOCUTIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: AND USHS THEREOF
FILE REFERENCE: CL001172
CURRENT APPLICATION NUMBER: US/09/813,320
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 397658
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(397658)
OTHER INFORMATION: n = A,T,C or G
US-09-813-320-3
```

```
Alignment Scores:
Pred. No.: 9.57e+04 length: 397658
Score: 35.00 Matches: 6
Percent Similarity: 90.91% Conservative: 4
Best Local Similarity: 54.55% Mismatches: 1
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0
```

US-09-823-649a-6 (1-11) x US-09-813-320-3 (1-397658)

```
OY 1 LeuSerlysaRg1leGlyleuSerValSerGlu 11
||| |||||:||||| ||| |||
Db 101718 CTCCTAAGAGATGGCGATCATGATTTTCAGAA 101750
```

RESULT 13

```
US-09-954-456-1202
Sequence 1202, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
```

```
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1202
LENGTH: 431
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1202
```

```
Alignment Scores:
Pred. No.: 50.2 length: 431
Score: 34.00 Matches: 6
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 60.00% Mismatches: 0
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0
```

US-09-823-649a-6 (1-11) x US-09-954-456-1202 (1-431)

```
OY 1 LeuSerlysaRg1leGlyleuSerValSer 10
||| |||||:||||| ||| |||
Db 263 GTCCTAAGAGAGGTGGCGCTTCACATCAGT 292
```

```
RESULT 14
US-09-880-107-709
Sequence 709, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
TITLE OF INVENTION: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 709
LENGTH: 431
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA292440
US-09-880-107-709

Alignment Scores:
Pred. No.: 50.2 Length: 431
Score: 34.00 Matches: 6
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 60.00% Mismatches: 0
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-6 (1-11) x US-09-880-107-709 (1-431)

QY 1 LeuSerLysArgIleGlyLeuSerValSer 10
      ::::::::::::::::::::::::::::
Db 263 GTCTCAGGAGGCGTGGCTTCTGTCAGT 292

RESULT 15
US-09-864-761-5243
; Sequence 5243, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5243

```

```

; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004954.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
US-09-864-761-5243

Alignment Scores:
Pred. No.: 54.5 Length: 462
Score: 34.00 Matches: 7
Percent Similarity: 88.89% Conservative: 1
Best Local Similarity: 77.78% Mismatches: 0
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-6 (1-11) x US-09-864-761-5243 (1-462)

QY 1 LeuSerLysArgIleGlyLeuSerVal 9
      ||| ||::::::::::::::::::::
Db 187 TTACAGAAAAGCGCGCTTCTCTGTC 213

```

Search completed: January 21, 2003, 10:08:27
 Job time : 66.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds

(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649a-6

Perfect score: 49

Sequence: 1 LSKRIGLSVSE 11

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -
-MODEL=frame_plus_p2n_model -DEV=xip
-O/corn2.1/USPTO.spool/US09823649/runat.21012003.093151.24616/app.query.fasta.1.1393
-DB=EST -QFMT=faststep -SUFFIX=est -MINMATCH=0.1 -LOOPC=0 -DOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09823649 @CGN.1.1.2820 @runat.21012003.093151.24616 -NCPU=6 -ICPU=3
-NO.XIPXY -NO.MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

EST:
1: em_estda:
2: em_estnum:
3: em_estin:
4: em_estm:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_hlc:
9: gb_est1:
10: gb_est2:
11: gb_hlc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: gb_gss:
18: em_gss_hum:
19: em_gss_inv:
20: em_gss_pln:
21: em_gss_vrt:
22: em_gss_fun:
23: em_gss_man:
24: em_gss_mus:
25: em_gss_other:
26: em_gss_pro:
27: em_gss_rtd:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
C 1	41	83.7	429	9 AU228362	AU228362 AU228362
C 2	40	81.6	509	17 AU2929376	AU2929376 479.dlf25
C 3	40	81.6	1117	17 C806044N	Al405021 T7 end of
C 4	39	79.6	383	10 AM669477	AM669477 MR3-TS006
C 5	39	79.6	691	17 AG085155	AC085155 Pan tsm01
C 6	38	77.6	413	12 BG096912	BG096912 EST641431
C 7	38	77.6	425	14 BM671118	BM671118 M950133A
C 8	38	77.6	442	10 AM057231	AM057231 T81405.Y
C 9	38	77.6	495	17 BM628659	BM628659 BACp28-C
C 10	38	77.6	545	12 BF049437	BF049437 db81d03.X
C 11	38	77.6	613	10 AM958920	AM958920 EST370990
C 12	38	77.6	666	13 BJ165314	BJ165314 BJ165314
C 13	38	77.6	812	14 B0506021	B0506021 EST613436
C 14	37	75.5	313	9 AA572148	AA572148 v150e12.f
C 15	37	75.5	350	9 AU112697	AU112697 AU112697
C 16	37	75.5	368	13 BJ389576	BJ389576 BJ389576
C 17	37	75.5	376	14 C10047	C10047 C10047 Yuf1
C 18	37	75.5	378	9 AV188048	AV188048 AV188048
C 19	37	75.5	413	17 B71747	B71747 T31847F T4M
C 20	37	75.5	487	13 BJ325691	BJ325691 BJ325691
C 21	37	75.5	497	12 BG018916	BG018916 db802804.
C 22	37	75.5	502	13 BJ362913	BJ362913 BJ362913
C 23	37	75.5	515	17 AQ788974	AQ788974 HS_3044.B
C 24	37	75.5	526	14 BM953398	BM953398 952061H02
C 25	37	75.5	541	12 AZ258482	AZ258482 PCCT-23-1
C 26	37	75.5	541	12 BF246262	BF246262 601853864
C 27	37	75.5	554	13 BJ363889	BJ363889 BJ363889
C 28	37	75.5	556	13 BJ410239	BJ410239 BJ410239
C 29	37	75.5	576	13 BJ093481	BJ093481 BJ093481
C 30	37	75.5	596	13 BJ142376	BJ142376 BJ142376
C 31	37	75.5	599	17 AZ130889	AZ130889 CSJNBB007
C 32	37	75.5	632	9 AL681825	AL681825
C 33	37	75.5	654	13 BJ070414	BJ070414 BJ070414
C 34	37	75.5	662	13 BJ075163	BJ075163 BJ075163
C 35	37	75.5	668	13 BJ346356	BJ346356 BJ346356
C 36	37	75.5	705	17 AG105753	AG105753 Pan t10q1
C 37	37	75.5	707	14 BJ0181293	BJ0181293 UI-n-EUO-
C 38	37	75.5	723	9 A1936015	A1936015 w97912.x
C 39	37	75.5	728	12 BG494983	BG494983 602541612
C 40	37	75.5	751	12 BF615513	BF615513 de08004.Y
C 41	37	75.5	760	13 B1903478	B1903478 603167007
C 42	37	75.5	779	17 BH305311	BH305311 CH230-192
C 43	37	75.5	797	17 BH094735	BH094735 PCCT-24-2
C 44	37	75.5	805	12 BG722042	BG722042 602698632
C 45	37	75.5	809	12 BF541083	BF541083 602068667

ALIGNMENTS

RESULT 1
AU228362/c 429 bp mRNA linear EST 23-APR-2002
DEFINITION AU228362 RAT16 Arabidopsis thaliana cDNA clone RAT16-13-107 3',
LOCUS mRNA sequence.
ACCESSION AU228362
VERSION AU228362
KEYWORDS GI:19743009
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 429)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,

TITLE Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinzaki, K.
JOURNAL Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msek@rcc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda PhiC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further details.

FEATURES**source**

1. 429
 /location/Qualifiers
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone_1fb="RAP16-13-107"
 /clone_1lb="RAP16"
 /lab_host="DH10B"
 /note="Site_1: BamHI; Site_2: SalI; dark-grown"

BASE COUNT 156 a 86 c 74 g 113 t
ORIGIN

Alignment Scores:

Pred. No.: 153
Score: 41.00
Percent Similarity: 90.91%
Best Local Similarity: 81.82%
Query Match: 83.67%
DB: 9
Gaps: 0

US-09-823-649A-6 (1-11) x AU228362 (1-429)

QY 1 LeuSerLyArgIleGlyLeuSerValSerGlu 11
 |||||
Db 123 TTTCACAAACATCGTTTGAAGGTTTCACAA 91

RESULT 2

AZ929376/c

LOCUS AZ929376 509 bp DNA linear GSS 01-APR-2001
DEFINITION 479.dif2se12.s1 Saccharomyces kluyveri Saccharomyces kluyveri
ACCESSION AZ929376
VERSION AZ929376.1 GI:13500283
KEYWORDS GSS.
SOURCE Saccharomyces kluyveri.
ORGANISM Saccharomyces kluyveri.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE

AUTHORS

TITLE**JOURNAL****COMMENT**

1 (bases 1 to 509)
 Clifton, P.F., Hillier, L.W., Fulton, L., Graves, T., Miner, T., Gish
 , W.R., Waterston, R.H., and Johnston, M.
 Surveying Saccharomyces genomes to identify functional elements by
 comparative DNA sequence analysis
 Unpublished (2001)
 Contact: Johnston M
 Department of Genetics
 Washington University Medical School
 Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
 Tel: 314 362 2735
 Fax: 314 362 7855
 Email: mj@genetics.wustl.edu
 Class: random plasmid subclone.
 location/Qualifiers

FEATURES**source**

1. 509
 /organism="Saccharomyces kluyveri"

/strain="NRRL Y-12651 (CBS 3082)"
 /db_xref="taxon:4934"
 /clone_1fb="479.dif2se12.s1"
 /clone_1lb="Saccharomyces kluyveri"
 /note="Random genomic sequence"

BASE COUNT

154 a 103 c 84 g 168 t

ORIGIN

154 a 103 c 84 g 168 t

Alignment Scores:

Pred. No.: 309
Score: 40.00
Percent Similarity: 81.82%
Best Local Similarity: 81.82%
Query Match: 81.63%
DB: 17
Gaps: 0

US-09-823-649A-6 (1-11) x AZ929376 (1-509)

QY 1 LeuSerLyArgIleGlyLeuSerValSerGlu 11
 |||||
Db 101 TTGTCAAAGATTGGCTACTGTAAAGAG 69

RESULT 3

CNS06M4N/c

LOCUS CNS06M4N 1117 bp DNA linear GSS 17-JUN-2001
DEFINITION 77 end of clone AU0AA002A02 of library AU0AA from strain CBS 3082
ACCESSION AU05021
VERSION AU05021.1 GI:12167229
KEYWORDS GSS.
SOURCE Saccharomyces kluyveri.
ORGANISM Saccharomyces kluyveri.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE

AUTHORS

1 (bases 1 to 1117)
 Soucier, J.L., Algie, M., Attiguenave, F., Blandin, G.,
 Boitot, F., Boudry, M., Bon, E., Brothier, P., Casaregola, S.,
 de-Montigny, J., Dujon, B., Durieux, P., Lepingle, A., Llorente, B.,
 Maupertuy, A., Neuvéglise, C., Ozler, K., Kalogeropoulos, O., Peller, S.,
 Saurin, W., Tekala, F., Toffano-Nicolas, C., Wesolowski-Louvel, M.,
 Wincker, P., and Weissenbach, J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies

JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876

REFERENCE

AUTHORS

TITLE**JOURNAL****COMMENT**

2 (bases 1 to 1117)
 Neuvéglise, C., Bon, E., Lepingle, A., Wincker, P., Attiguenave, F.,
 Gattilard, J., and Casaregola, S.
 Genomic exploration of the hemiascomycetous yeasts: 9.
 Saccharomyces kluyveri
 FEMS Lett. 487 (1), 56-60 (2000)
PUBMED 20584719
MEDLINE 11152884
PUBMED 11152884
 3 (bases 1 to 1117)
 Genoscope.
 Direct Submission
 Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marianus var. marianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES**source**

1. 1117
 /organism="Saccharomyces kluyveri"

```

misc-feature
<31..>1042
/note="similar to Saccharomyces cerevisiae ORF YHR099w [
TRAI ; strong similarity to human TRRAP protein ]
1 putative frameshift(s)"
/evidence=not experimental

BASE COUNT      332 a      230 c      186 g      367 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      889      Length:      1117
Score:          40.00      Matches:      9
Percent Simiarity: 81.82%      Conservative: 0
Best Local Similarity: 81.82%      Mismatches:  2
Query Match:      81.63%      Indels:      0
DB:              17      Gaps:      0

US-09-823-649a-6 (1-11) x CNS06MAN (1-1117)

QY      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db      482 TTGTCAAAAGATGGCTACGTGTAAGAGAG 450

RESULT 4
AM869477/c      383 bp      mRNA      linear      EST 22-MAY-2000
LOCUS
DEFINITION      MR3-SN0068-070500-002-b10 SN0068 Homo sapiens cDNA, mRNA sequence.
ACCESSION      AM869477
VERSION      AM869477.1      GI:8003530
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE      1 (bases 1 to 383)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., DeOliveira,P.S., Bucher,P., Joneneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202683
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/genhtml2.pl?l1=612-MR3-SN0068-070
500-002-b10&l3=2000-05-07&l4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 197.
Location/Qualifiers
1..383
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="SN0068"
/dev_stage="Adult"
/note="Organ: stomach; normal; Vector: puc18; Site:1: Smar;
Site 2: Smar; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application

```

```

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      134 a      102 c      87 g      60 t
ORIGIN

Alignment Scores:
Pred. No.:      339      Length:      383
Score:          39.00      Matches:      8
Percent Similarity: 100.00%      Conservative:  3
Best Local Similarity: 72.73%      Mismatches:  0
Query Match:      79.59%      Indels:      0
DB:              10      Gaps:      0

US-09-823-649a-6 (1-11) x AM869477 (1-383)

QY      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db      140 CTGTCAGAGAAATGGCTTACGTGACGAG 108

RESULT 5
AG085155
AG085155
LOCUS
DEFINITION      Pan troglodytes DNA, clone: PTB-083B22.F, genomic survey sequence.
ACCESSION      AG085155
VERSION      AG085155.1      GI:16636957
KEYWORDS      GSS.
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC library clone:PTB-083B22.F.
ORGANISM      Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE      1
AUTHORS      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tokoki,Y., Matanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of library PTB
JOURNAL      Unpublished
2 (bases 1 to 691)
AUTHORS      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tokoki,Y., Matanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL      Submitted (02-NOV-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho, Wsuri-mi-ku, Yokohama, Kanagawa 230 0045, Japan
(E-mail:chimbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT      Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the Rad process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI.
Location/Qualifiers
1..691
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-083B22.F"
/sex="Male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC library"

BASE COUNT      230 a      151 c      127 g      181 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      750      Length:      691
Score:          39.00      Matches:      8
Percent Similarity: 90.91%      Conservative:  2
Best Local Similarity: 72.73%      Mismatches:  1
Query Match:      79.59%      Indels:      0

```


AUTHORS

McCarter, J., Clifton, S., Chiappelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Merra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Stepien, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

TITLE

The Washington Univ. Nematode EST Project, 1999

JOURNAL

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Dr. Ralf Sommer DNA Sequencing by:

Washington University Genome Sequencing Center

Contact Dr. Ralf Sommer (ralf.sommer@wustl.edu) for

information about this clone.

Putative full length read

The vector to vector length is 738

FEATURES

Seq primer: -40RP from Gibco

High quality sequence stop: 404.

Location/Qualifiers

1..442

/organism="Pristionchus pacificus"

/strain="PS 312"

/db_xref="taxon:54126"

/clone_lib="Sommer Pristionchus"

/sex="predominantly hermaphroditic"

/dev_stage="mixed stages (embryo to adult)"

/lab_host="not applicable (host cell line)"

/note="Vector: uni-zap XR Vector (Stratagene); Site: 1: 5'

EcoRI; Site: 2: 3' XhoI; 1st strand cDNA was primed with a

XhoI - oligo(dT) primer. Double stranded cDNA was ligated

to EcoRI adaptors digested with XhoI and cloned into XhoI

and EcoRI sites. Primary complexity of the library was 10

in the 7th. The library went through one round of

amplification"

US-09-823-649A-6 (1-11) x AM097231 (1-442)

QY 1 LeuSerLySArgIleGlyLeuSerValSerGlu 11

|||||.....:|||||

DB 164 TTGAGCAGAGGCTCGAATTGACTTAAGTAA 196

RESULT 9

BH828859

LOCUS

DEFINITION

BACPP28-G18.Y Pristionchus pacificus BAC ends Pristionchus

pacificus genomic DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 495)

Srinivasan, J., Siz, W., Lanz, C., Brand, A., Nandakumar, R., Raddatz, G., Witte, H., Keller, H., Kipping, I., Pires da Silva, A., Jesse, T., Millare, J., de Both, M., Schuster, S.C. and Sommer, R.J.

TITLE

A BAC-based genetic linkage map of the nematode Pristionchus

pacificus

Unpublished (2002)

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

Class: BAC ends.

Location/Qualifiers

1..495

/organism="Pristionchus pacificus"

/strain="var. California"

/db_xref="taxon:54126"

/clone_lib="Pristionchus pacificus BAC ends"

BASE COUNT

142 a 85 c 105 g 163 t

ORIGIN

Alignment Scores:

Pred. No.: 771

Score: 38.00

Percent Similarity: 90.91%

Best Local Similarity: 63.64%

Query Match: 77.55%

DB: 17

Gaps: 0

US-09-823-649A-6 (1-11) x BH828859 (1-495)

QY 1 LeuSerLySArgIleGlyLeuSerValSerGlu 11

|||||.....:|||||

DB 269 TTGAGCAGAGGCTCGAATTGACTTAAGTAA 301

RESULT 10

BFO49437

LOCUS

DEFINITION

BFO49437

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 545)

Clifton, S., Johnson, S.L., Njumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

Unpublished (1999)

Contact: Sandy Clifton, Ph.D.

Washu Xenopus EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by N. Garrett, P. Lemaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute). DNA Sequencing by: Washington

University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available

through the I.M.A.G.E. Consortium/MLM at: info@image.lim.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 423.

Location/Qualifiers

1..545

/organism="Xenopus laevis"

US-09-823-649A-6 (1-11) x AM097231 (1-442)

QY 1 LeuSerLySArgIleGlyLeuSerValSerGlu 11

|||||.....:|||||

DB 269 TTGAGCAGAGGCTCGAATTGACTTAAGTAA 301

RESULT 10

BFO49437

LOCUS

DEFINITION

BFO49437

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 545)

Clifton, S., Johnson, S.L., Njumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

Unpublished (1999)

Contact: Sandy Clifton, Ph.D.

Washu Xenopus EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by N. Garrett, P. Lemaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute). DNA Sequencing by: Washington

University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available

through the I.M.A.G.E. Consortium/MLM at: info@image.lim.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 423.

Location/Qualifiers

1..545

/organism="Xenopus laevis"

```

/db_xref="taxon:8355"
/clone="IMAGE:3379397"
/clone_lib="Wellcome CRC psk egg"
/tissue_type="egg"
/lab_host="DH10B (phage-resistant)"
/notes="Vector: Bluescript SK-; Site_1: NotI; Site_2:
ECORI; CDNAs were oligo-dT primed and directionally
cloned. Library was constructed by N. Garrett, P. Lemaire,
A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute).
Note: This is a Xenopus Gene Collection (XGC) library."

BASE COUNT      129 a      153 c      105 g      158 t
ORIGIN

Alignment Scores:
Pred. No.:      877      Length:      545
Score:          38.00      Matches:      8
Percent Similarity: 90.00%      Conservative: 1
Best Local Similarity: 80.00%      Mismatches: 1
Query Match:     77.55%      Indels:      0
DB:              12      Gaps:          0

US-09-823-649a-6 (1-11) x BF049437 (1-545)

QY      1 Leuserlysrq1leqlyeuserValser1u 10
||||| ||||||||| |||||
Db      508 CTCAGCTCAGCATAGGTTATCTATCTCA 537

RESULT 11
AM958920/c      613 bp      mRNA      linear      EST 01-JUN-2000
LOCUS      AM958920
DEFINITION      EST370990 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION      AM958920
VERSION      AM958920.1
KEYWORDS      GI:8148604
SOURCE      EST.
ORGANISM      human.
REFERENCE      Human.
AUTHORS      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
              Hegde, P., Qi, R., Abernathy, K., Dharrap, S., Gaspard, R., Gay, C., Holt
              1, E., Saeed, A. I., Sharov, V., Lee, N. H., Yestman, T. J. and
              Quackenbush, J.
              Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
              Unpublished (2000)
              Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 127
              Seq primer: Reverse.

FEATURES
Source      Location/Qualifiers
1..613
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/notes="Vector: BluescriptSKm"

BASE COUNT      175 a      130 c      158 g      149 t
ORIGIN

Alignment Scores:
Pred. No.:      1,03e+03      Length:      613
Score:          38.00      Matches:      8
Percent Similarity: 90.91%      Conservative: 2
Best Local Similarity: 72.73%      Mismatches: 1
Query Match:     77.55%      Indels:      0
DB:              10      Gaps:          0

US-09-823-649a-6 (1-11) x AM958920 (1-613)

QY      1 Leuserlysrq1leqlyeuserValser1u 11
||||| ||||||||| |||||

```

```

||||| ||||||||| |||||
Db      606 CTTCAAGAGATGGCTTAAGCTTCAGAA 574

RESULT 12
B0165314/c      666 bp      mRNA      linear      EST 24-JAN-2002
LOCUS      B0165314
DEFINITION      B0165314 full length cDNA library, chloronemata and young
              gametophores Physcomitrella patens subsp. patens cDNA clone pp13n09
              3', mRNA sequence.
ACCESSION      B0165314
VERSION      B0165314.1
KEYWORDS      GI:18333300
SOURCE      EST.
ORGANISM      Physcomitrella patens subsp. patens.
              Physcomitrella patens subsp. patens.
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
              Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
              1 (bases 1 to 666)
              Fujita, T., Shin, I. T., Seki, M., Kamiya, A., Uchiyama, T., Nishiyama, T.,
              Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe
              M.
              Comparison of the moss Physcomitrella patens genome with flowering
              plants genome
              Unpublished (2002)
              Contact: Tadao Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshin@genes.nig.ac.jp
              A backbone of the vector is basically from pbluescript(KS), that
              was in vivo excised from a modified IPS phage vector (Mo bi Tec,
              Germany). 5' end of the cDNA that was digested with XhoI was
              ligated to SalI site of the vector and the 3' end including polyA
              tail was ligated to BamHI site of the vector. cDNA insert could be
              amplified with conventional T7 and T3 primers. This full-length
              cDNA library was generated basically according to the method
              described in The Plant J 15, 707-720 (1998) Seki M. et al.
              Protonemata were blended by the POLYPRON, and then cultivated on
              the BCDM2 medium for 13-14 days under the continuous light.

FEATURES
Source      Location/Qualifiers
1..666
/organism="Physcomitrella patens subsp. patens"
/db_xref="taxon:145481"
/clone="Pp13n09"
/clone_lib="full length cDNA library, chloronemata and
              young gametophores"
/tissue_type="mixture of chloronemata and young
              gametophores with 2 to 5 leaves"

BASE COUNT      150 a      208 c      141 g      167 t
ORIGIN

Alignment Scores:
Pred. No.:      1.15e+03      Length:      666
Score:          38.00      Matches:      9
Percent Similarity: 90.91%      Conservative: 1
Best Local Similarity: 81.82%      Mismatches: 1
Query Match:     77.55%      Indels:      0
DB:              13      Gaps:          0

US-09-823-649a-6 (1-11) x B0165314 (1-666)

QY      1 Leuserlysrq1leqlyeuserValser1u 11
||||| ||||||||| |||||
Db      94 TTGCTCGGAGANTGGCTATGAGTCTCGGAA 62

RESULT 13
B0506021/c      812 bp      mRNA      linear      EST 22-JUL-2002
LOCUS      B0506021
DEFINITION      EST613436 Generation of a set of potato cDNA clones for microarray
              analyses mixed potato tissues Solanum tuberosum cDNA clone SWMGH75
              3' end, mRNA sequence.
ACCESSION      B0506021

```


source

1. .350
 /organism="Caenorhabditis elegans"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone_lib="K745h9"
 /clone_lib="unpublished oligo-capped cDNA library"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="varied"
 /note="Organ: ovary; Vector: lambd zap II; Approximately
 1,000 ovaries were hand dissected from adult C. elegans.
 RNA was extracted and the cDNA was amplified using the
 SMART PCR cDNA Library Construction Kit (Clontech). This
 was ligated into Lambda Zap."

BASE COUNT 95 a 79 c 86 g 90 t
 ORIGIN

Alignment Scores:

Pred. No.:	778	Length:	350
Score:	37.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	4
Best Local Similarity:	63.64%	Mismatches:	0
Query Match:	75.51%	Indels:	0
DB:	9	Gaps:	0

US-09-823-649a-6 (1-11) x AU112697 (1-350)

QY 1 leuserlyargilegyleuserValserciu 11

Db 212 ATCAACAGAGAGATTGCTATCATTCACAA 180

Search completed: January, 21, 2003, 12:24:26
 Job time : 1172.14 secs

US-09-823-649A-5 (1-11) x US-09-815-242-9415 (1-2031)

QY 1 LeuSeValArgLeuGly***ProValysGu 11
|||||
Db 70 TTATCTGTTAAGTGAGCTTACTATTAAGGAA 102

RESULT 13

US-09-286-240-5/c
; Sequence 5, Application US/09286240
; Patent No. US20020010320A1
; GENERAL INFORMATION:
; APPLICANT: Felt, James W
; TITLE OF INVENTION: Chimeric and Humanized Antibodies to Angiogenin
; FILE REFERENCE: 10498/74073
; CURRENT APPLICATION NUMBER: US/09/286,240
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1809)..(2252)
US-09-286-240-5

Alignment Scores:

Pred. No.: 544 Length: 4668
Score: 34.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-5 (1-11) x US-09-286-240-5 (1-4668)

QY 1 LeuSeValArgLeuGly***Pro 8
|||||
Db 4038 CTTTCAGTCAGGCTGGCTCTCCT 4015

RESULT 14

US-09-863-777-1/c
; Sequence 1, Application US/09863777
; Patent No. US20020019359A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Karen A.
; TITLE OF INVENTION: Antisense Inhibition of Angiogenin Expression
; FILE REFERENCE: 10498/05286
; CURRENT APPLICATION NUMBER: US/09/863,777
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/041182
; PRIOR FILING DATE: 1997-03-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1809)..(2252)
US-09-863-777-1

Alignment Scores:

Pred. No.: 544 Length: 4668
Score: 34.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-5 (1-11) x US-09-863-777-1 (1-4668)

QY 1 LeuSeValArgLeuGly***Pro 8
|||||
Db 4038 CTTTCAGTCAGGCTGGCTCTCCT 4015

RESULT 15

US-09-880-107-2239/c
; Sequence 2239, Application US/09880107
; Patent No. US/0020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2239
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M11567
US-09-880-107-2239

Alignment Scores:

Pred. No.: 544 Length: 4668
Score: 34.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649A-5 (1-11) x US-09-880-107-2239 (1-4668)

QY 1 LeuSeValArgLeuGly***Pro 8
|||||
Db 4038 CTTTCAGTCAGGCTGGCTCTCCT 4015

Search completed: January 21, 2003, 10:07:50
Job time: 33.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 21, 2003, 09:31:57 (Search time 1168.14 seconds)

(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649a-5
Perfect score: 49
Sequence: 1 LSVRLGXPKYE 11

Scoring table:

BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Command line parameters: -
-MODEL=frame_plus.p2n model -DRV=1p
-Q/cgn2.1/USPFO.spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB-EST-QFWT=fastac -SUFFIX=1st -MINMATCH=0.1 -DOPEL=0 -DOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -M <ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0982364_@CGN.1.1.2820 @runat.21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAB -LARGEOTHER -NEG_SCORES=0 -MATT -IONGLOC -DRV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -DELop=6 -DELEXt=7

Database :

RST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_estl:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_estfun:*
15: em_estfun:*
16: em_estlom:*
17: gb_gsa:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	85.7	589	17 BH110033	BH110033 RPCI-24-3
2	41	83.7	584	14 BH008878	BH008878 OCH6011.Y
3	41	83.7	985	12 BE784762	BE784762 601473679
4	40	81.6	350	10 AM047855	AM047855 UT-M-BH1
5	40	81.6	529	12 BS102214	BS102214 RH122_22
6	39	79.6	649	17 A7795132	A7795132 280049912
7	39	79.6	658	12 BP214447	BP214447 601845635
8	39	79.6	720	17 BH699492	BH699492 BOMBK74TF
9	38	77.6	112	9 AA293830	AA293830 z165B09.F
10	38	77.6	190	14 B0565397	B0565397 q136g12.Y
11	38	77.6	259	17 A2769101	A2769101 180569324
12	38	77.6	300	13 B1690928	B1690928 603312076
13	38	77.6	309	17 BH743626	BH743626 q125a07.B
14	38	77.6	338	10 AM477975	AM477975 17923.MAR
15	38	77.6	384	17 A2465796	A2465796 IM0275N19
16	38	77.6	405	17 CNS05R1K	AL349985 Telradon
17	38	77.6	415	10 AM772450	AM772450 hn74c08.X
18	38	77.6	432	9 AA337681	AA337681 vk44f02.F
19	38	77.6	517	10 AM785618	AM785618 116849.MA
20	38	77.6	523	13 B1507078	B1507078 BM170023B
21	38	77.6	531	9 A1596792	A1596792 vk44f02.Y
22	38	77.6	546	10 AM642017	AM642017 cm14a06.W
23	38	77.6	632	17 A2460045	A2460045 IM0265A18
24	38	77.6	638	13 B1186406	B1186406 UNI-P-FN-
25	38	77.6	674	17 AC064313	AC064313 Pan.trog1
26	38	77.6	685	12 BP702429	BP702429 MI-P-A2.a
27	38	77.6	699	13 B1184838	B1184838 OMD-P-A2.a
28	38	77.6	716	17 AG076995	AG076995 Pan.trog1
29	37	75.5	310	9 AV154670	AV154670 AV154670
30	37	75.5	326	12 H0661863	H0661863 Kx07c12.Y
31	37	75.5	336	13 B1344292	B1344292 372728.MA
32	37	75.5	371	17 A7319685	A7319685 1M0039G21
33	37	75.5	417	17 A0338684	A0338684 IRS-2210.B
34	37	75.5	450	10 AM387729	AM387729 MR-ST011
35	37	75.5	468	12 BF387658	BF387658 UI-R-CA1
36	37	75.5	475	9 AA564641	AA564641 n125e12.s
37	37	75.5	496	10 A5W581884	A5W581884 MR4-ST011
38	37	75.5	510	9 A1562393	A1562393 TENS2383
39	37	75.5	533	14 BM884318	BM884318 IC211003.Y
40	37	75.5	640	12 H033871	H033871 601456259
41	37	75.5	691	17 BH722607	BH722607 BOMCLO2TF
42	37	75.5	742	14 H0573626	H0573626 UI-M-FD0-
43	37	75.5	979	17 CNS02E1J	AL193780 Tetradon
44	37	75.5	1004	12 H023266	H023266 RS132872
45	37	75.5	1052	13 BM548113	BM548113 ASENCOURT

ALIGNMENTS

RESULT 1
BH110033
LOCUS
DEFINITION
RPCI-24-323K23.TJ RPCI-24 Mus musculus genomic clone RPCI-24-323K23
, DNA sequence.
ACCESSION
BH110033
VERSION
BH110033.1 GI:14943718
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Fukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 589)
REFERENCE
Zhao,S., Niernan,W., Malek,J., Shatsman,S., Akintiret,B., Levins,M.,
Tsaye,G., Geer,K., Krol,M., Shwartsbeyn,A., Gebregeorgis,E.,

TITLE
JOURNAL
COMMENT

Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-323K23.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@igrr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdcjong@mail.cho.org). Clones may be purchased from BACPAC
Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 323 row: K column: 23
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..589
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-323K23"
/clone_id="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: PTARBAC1; Site 1: BamHI; Site 2: BamHI;
RPCI-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the PTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 157 a 137 c 170 g 125 t
ORIGIN
Alignment Scores:
Pred. No.: 534 Length: 589
Score: 42.00 Matches: 8
Percent Similarity: 81.82% Conserved: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 85.71% Indels: 0
DB: 17 Gaps: 0

US-09-823-649a-5 (1-11) x BH110033 (1-589)

OY 1 LeuserValArgleugly***ProValys61u 11
|||||
DB 61 CTATCAGTCAAGCTCGGAGAGCCCATCAGTCAG 93

RESULT 2
BU008878 584 bp mRNA linear EST 22-AUG-2002
LOCUS OGH8011.yg.abi OG_FRGH lettuce serritola lactuca sativa cDNA clone
DEFINITION OGH8011, mRNA sequence.
ACCESSION BU008878
VERSION BU008878.1 GI:22443285
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 584)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, W., Lavelle, D., Chevalier, P., Ziegler, J., Ellison,
P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://comgenomics.ucdavis.edu/>
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]

FEATURES
source

Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@ucdavis.edu]
Belongs to contig OG_CA_Contig932, see <http://cgpd.ucdavis.edu/>
for details.
Plate: OGH8 row: O column: 11.
Location/Qualifiers
1..584
/organism="Lactuca sativa"
/cultivar="L.serritola"
/db_xref="taxon:4236"
/clone="OGH8011"
/clone_id="OG_FRGH lettuce serritola"
/lab_host="E.coli"
/note="Vector: pBRC/NASFIAB. The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at <http://cgpd.ucdavis.edu/>
TMS_SEQ-Not found"

BASE COUNT 170 a 123 c 131 g 160 t
ORIGIN
Alignment Scores:
Pred. No.: 804 Length: 584
Score: 41.00 Matches: 9
Percent Similarity: 81.82% Conserved: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 83.67% Indels: 0
DB: 14 Gaps: 0

US-09-823-649a-5 (1-11) x BU008878 (1-584)

OY 1 LeuserValArgleugly***ProValys61u 11
|||||
DB 213 CTATCTTAAGCCAGCGGACCGCTGAAAGAC 181

RESULT 3
BE784762 985 bp mRNA linear EST 20-OCT-2000
LOCUS BE784762 60147867941 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:3876691 5',
DEFINITION mRNA sequence.
ACCESSION BE784762
VERSION BE784762.1 GI:10205960
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 985)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gaardar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: FLAM9637 row: n column: 20
High quality sequence stop: 287.
Location/Qualifiers
1..985

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3876691"
 /clone.lib="NIH_MGC_68"
 /tissue.type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.8 kb. library constructed by Life
 Technologies."
 BASE COUNT 290 a 312 c 254 g 129 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.34e+03 Length: 985
 Score: 41.00 Matches: 7
 Percent Similarity: 81.82% Conservative: 2
 Best Local Similarity: 63.64% Mismatches: 2
 Query Match: 83.67% Indels: 0
 DB: 12 Gaps: 0

US-09-823-649A-5 (1-11) x BE784762 (1-985)

QY 1 LeuserValArgLeugly**ProVallysglu II
 |||||:||||:|||||
 Db 283 CTCAGCATCGCTATTGGCACACCAAGAG 315

RESULT 4

AM047855

LOCUS AM047855 350 bp mRNA linear EST 18-SEP-1999
 DEFINITION UI-M-BH1-als-e-04-0-UI-s1 NIH_BMAP_M_S2 Mus musculus cDNA clone
 UI-M-BH1-als-e-04-0-UI 3', mRNA sequence.

ACCESSION AM047855

VERSION AM047855.1 GI:5908384

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
 1 (bases 1 to 350)

REFERENCE Ronald M.F., Lennon G. and Soares M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT

National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mestr@mail.nih.gov
 Oligo-dT track not found. Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
 available by the means that is soon to be determined. When NIH
 determines the means for distribution of the BMAP cDNA clones, this
 record will be updated accordingly when that means is determined.
 Seq primer: M13 Forward
 POLYA-No.

FEATURES

source

1..350 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH1-als-e-04-0-UI"
 /clone.lib="NIH_BMAP_M_S2"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_M_S2 library is a subtracted library derived from
 NIH_BMAP_M_S1, which in turn is a subtracted library

derived from a mixture of normalized libraries from ten
 regions of the mouse brain (cerebellum, brain stems,
 olfactory bulbs, hypothalamus, cortex, amygdala, basal
 ganglia, pineal gland, striatum, hippocampus). The driver
 used for subtraction consisted of a pool of 5,000 clones
 from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
 obtained from non-normalized and normalized mouse brain
 spinal cord libraries.
 TAG_Lib="NIH_BMAP_M_S2"
 TAG_Tissue="Prefrontal-cortex"
 TAG_SKO="CCFCA"
 BASE COUNT 76 a 107 c 98 g 69 t
 ORIGIN

Alignment Scores:

Pred. No.: 737 Length: 350
 Score: 40.00 Matches: 8
 Percent Similarity: 81.82% Conservative: 1
 Best Local Similarity: 72.73% Mismatches: 2
 Query Match: 81.63% Indels: 0
 DB: 10 Gaps: 0

US-09-823-649A-5 (1-11) x AM047855 (1-350)

QY 1 LeuserValArgLeugly**ProVallysglu II
 |||||:||||:|||||
 Db 234 CTACCTCTTCGACAAAGCATACCCACTTAAGAA 266

RESULT 5

BG102214

LOCUS BG102214 529 bp mRNA linear EST 30-JAN-2001
 DEFINITION RH122_22.B12.B1_A003 Rhizome2 (RH122) Sorghum propinquum cDNA, mRNA
 sequence.

ACCESSION BG102214

VERSION BG102214.1 GI:12617047

KEYWORDS EST.

SOURCE Sorghum propinquum.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 529)
 Cordomier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt,
 L.H.

TITLE An EST database from Sorghum: Sorghum propinquum rhizomes

JOURNAL Unpublished (2000)

COMMENT Contact: Cordomier-Pratt MM

Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, 10n. 2502, Athens, GA 30602 7271, USA
 Tel: 706 542 1860
 Fax: 706 585 0210
 Email: mpratt@uga.edu
 Seq primer: JEN REV
 High quality sequence stop: 499
 POLYA-No.

FEATURES

source

1..529 Location/Qualifiers
 /organism="Sorghum propinquum"
 /db_xref="taxon:132711"
 /clone.lib="Rhizome2 (RH122)"
 /note="Organ: Rhizomes; Vector: pBluescript II from Lambda
 Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
 from poly-A RNA in the cloning vector lambda Zap II.
 Clones to be sequenced were prepared by mass excision."
 BASE COUNT 78 a 162 c 180 g 109 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.11e+03 Length: 529
 Score: 40.00 Matches: 8
 Percent Similarity: 81.82% Conservative: 1
 Best Local Similarity: 72.73% Mismatches: 2

QY 1 LeuserValArgLeugly***ProValLysGlu 11
 DB 135 ATGAGCCCTGAGATCGGAGGCCCGCCGACAGAG 416

RESULT 8
 BH699492/c BH699492 720 bp DNA linear GSS 20-FEB-2002
 LOCUS BOMK74TF_BO_2_3_KB Brassica oleracea genomic clone BOMK74, DNA
 DEFINITION sequence.
 ACCESSION BH699492
 VERSION BH699492.1 GI:18774833
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids 11; Brassicales; Brassicaceae; Brassica.

REFERENCE
 1 (bases 1 to 720)
 Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other GSSs: BOMK74TR
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: ctown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 1..720
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOMK74"
 /clone_1lb="BO_2_3_KB"
 /note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 164 a 239 c 147 g 170 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.27e+03 Length: 720
 Score: 39.00 Matches: 8
 Percent Similarity: 90.00% Conservative: 1
 Best Local Similarity: 80.00% Mismatches: 1
 Query Match: 79.59% Indels: 0
 Gaps: 0

US-09-823-649a-5 (1-11) x BH699492 (1-720)

QY 1 LeuserValArgLeugly***ProValLys 10
 DB 135 GTGAGGTGAGACTCGGTTCCCGGTGAG 106

RESULT 9
 AA293830/c 112 bp mRNA linear EST 12-AUG-1997
 LOCUS 2165309.r1 Soares_test15_NH Homo sapiens cDNA clone IMAGE:727193
 DEFINITION 5' similar to SW:MH1_HUMAN P40692 MOTT PROTEIN HOMOLOG 1; mRNA
 sequence.
 ACCESSION AA293830
 VERSION AA293830.1 GI:1941753
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1 (bases 1 to 112)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.

TITLE
 JOURNAL
 COMMENT

Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,
 T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through J.M.N.; contact the
 IMAGE Consortium (info@image.jml.gov) for further information.
 Trace considered overall poor quality
 Possible reversed clone; similarity on wrong strand
 Insert length: 645 Std Error: 0.00
 Seq primer: -28mb3 rev2 RT from Amersham
 High quality sequence stop: 1.

FEATURES
 source
 1..112
 /organism="Homo sapiens"
 /db_xref="GDB:5924105"
 /db_xref="taxon:9606"
 /clone="IMAGE:727193"
 /clone_1lb="Soares_test15_NH"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pTZ193D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Kco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech laboratories
 , Inc., and primed with a Not I - oligo(dT) primer 15'
 TGTTCACATCTGATGAGGAGCCGCCCGCCCATTTTCTTTTCTTTT 31'.
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Kco RI sites of the modified pTZ193 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bertio Soares and M. Fatima Bonaldo."

BASE COUNT 28 a 24 c 29 g 31 t
 ORIGIN

Alignment Scores:
 Pred. No.: 553 Length: 112
 Score: 38.00 Matches: 6
 Percent Similarity: 90.00% Conservative: 3
 Best Local Similarity: 60.00% Mismatches: 1
 Query Match: 77.55% Indels: 0
 Gaps: 0

US-09-823-649a-5 (1-11) x AA293830 (1-112)

QY 1 LeuserValArgLeugly***ProValLys 10
 DB 39 TTCTGATCAGATCGGATGCAATCAATCAAC 10

RESULT 10
 B0565397/c 190 bp mRNA linear EST 19-JUN-2002
 LOCUS g136g12.y1 Mouse Organ of Corti cDNA Phnucript Mus musculus cDNA
 DEFINITION clone g136g12 5', mRNA sequence.
 ACCESSION B0565397
 VERSION B0565397.1 GI:21468714
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 190)
 Kachar, B.
 TITLE EST analysis of gene expression in the mouse Organ of Corti at the
 onset of hearing
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kachar, B.
 Structural Cell Biology
 National Institute of Deafness and other Communication Disorders

50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
 Tel: 301-402-1599
 Fax: 301-402-1765
 Email: kacharbenid@nih.gov
 Plate: 36 row: g column: 12
 Seq primer: M13p1 reverse primer (ABI).
 Location/Qualifiers
 1. 190
 /organism="Mus musculus"
 /strain="BA1B/c"
 /db_xref="taxon:10090"
 /clone="g136912"
 /clone_lib="Mouse Organ of Corti cDNA pBluescript"
 /sex="male and female"
 /dev_stage="post natal day 5 to 13"
 /note="Organ: Organ of Corti; Vector: pBluescript; The organ of Corti (OC) was fine dissected from a total of 386 OC as follows: 102 samples from post-natal (P) day 5; 72 from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10; 14 from P12 and 24 from P13. After killing animals by cervical dislocation followed by decapitation, the bulla was removed and opened in Leibowitz medium. The bony capsule of the cochlea was chipped away, stria vascularis and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the micro FastTrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-Zap XR vector kit (catalog # 237211, Stratagene) and Uni-Zap XR Gigapack III Gold Cloning Kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer that contains an Xho I site. First strand synthesis was primed with the linker-primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adapters in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp, respectively. The cDNA was then directionally ligated to the Uni-Zap XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapack III Gold and, upon titration on XL1 Blue MRF' cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's ExAssist interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CACGACGACGACGAC) and 25x strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have known function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

BASE COUNT
 ORIGIN

45 a 57 c 48 g 40 t

Alignment Scores:
 Pred. No.: 930
 Score: 38.00
 Percent Similarity: 81.82%
 Best Local Similarity: 63.64%
 Query Match: 77.55%
 DB: 14
 Gaps: 0

US-09-823-649a-5 (1-11) x HQ565397 (1-190)

CY 1 LeuSERValAlaqlenugly***ProValLysGlu 11
 Db 146 TTGACGATCCGTTTGTCTGCTCTCTGTCAGACAG 114

RESULT 11
 A2769101/c 259 bp DNA linear GSS 16-FEB-2001
 LOCUS
 DEFINITION
 1M0569124F Mouse 10kb plasmid U06C1M library Mus musculus genomic
 clone U06C1M0569124 F. DNA sequence.
 ACCESSION
 A2769101 G11288888
 VERSION
 KEYWORDS
 GSS.

SOURCE
 OKAMISU
 house mouse.

REFERENCE
 AUTHORS
 Dunn,D., Aoyagi,A., Harber,M., Beacorn,T., Duval,B., Hamil,C.,
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 259)
 Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)

JOURNAL
 COMMENT
 Contact: Robert B. Weiss
 University of Utah
 University of Utah
 Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177

Email: dunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0569 row: j column: 24
 Seq primer: CGTGTAAACACGACGACGACG
 Class: plasmid ends
 High quality sequence stop: 259.

FEATURES
 Location/Qualifiers
 1..259
 /organism="Mus musculus"
 /strain="BA1B/c"
 /db_xref="taxon:10090"
 /clone="U06C1M0569124"
 /clone_lib="Mouse 10kb plasmid U06C1M library"
 /sex="male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: pMD22my; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD22 (g14732114/gb1aP129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to

adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 60 a 55 c 60 g 84 t

ORIGIN

Alignment Scores:

Pred. No.: 1.26e+03 Length: 259
Score: 38.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.72% Mismatches: 2
Query Match: 77.55% Indels: 0
DB: 17 Gaps: 0

US-09-823-649A-5 (1-11) x A2769101 (1-259)

QY 1 LeuserValArgLeugly**ProValysglu 11
|||||
Db 194 CTCGCTAAGACTCGAAGCCCGTGAAGAGC 162

RESULT 12

LOCUS B1690928 300 bp mRNA linear EST 18-SEP-2001
DEFINITION 603312076F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5352234 5',
mRNA sequence.
ACCESSION B1690928
VERSION B1690928.1 GI:15653557
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 300)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1895 row: 9 column: 19
High quality sequence stop: 300.
Location/Qualifiers

FEATURES

source

1..300
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5352234"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCW-Sport6; Site: 1; Salt: Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies, Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 60 a 91 c 75 g 74 t

ORIGIN

Alignment Scores:

Pred. No.: 1.46e+03 Length: 300
Score: 38.00 Matches: 6
Percent Similarity: 90.91% Conservative: 4
Best Local Similarity: 54.55% Mismatches: 1
Query Match: 77.55% Indels: 0
DB: 13 Gaps: 0

US-09-823-649A-5 (1-11) x B1690928 (1-300)

QY 1 LeuserValArgLeugly**ProValysglu 11
|||||
Db 41 CTTTCATCAGGCTAGGAGATCTTATCAGAG 73

RESULT 13

LOCUS BH743626 309 bp DNA linear GSS 25-FEB-2002
DEFINITION g125a07.b1 Bobuds01 Brassica oleracea genomic clone g125a07 5', DNA
sequence.
ACCESSION BH743626
VERSION BH743626.1 GI:18878240
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotys;
Rosidae; eustroids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

1 (bases 1 to 309)
Katarci, M., O'Shaughnessy, A., Palmer, L., Bahret, A., Baker, J., Ballia, V., Cummins, P.M., Katzenberger, F., King, L., Kirchhoff, K., Kull, K., Miller, B., Muller, S., Nascimento, L., Preston, R., Santos, L., Shan, R., Zlatavorn, I., Dechta, N., Rablinowicz, P.D., and McCombie, W.R.
Whole Genome Shotgun Reads from Brassica oleracea (2002b)
Unpublished (2002)
Contact: W. Richard McCombie
Lila Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshtl.org
Plate: g125 row: a column: 07
Seq primer: -21ntvFwd
Class: Shotgun
High quality sequence stop: 309.
Location/Qualifiers

FEATURES

source

1..309
/organism="Brassica oleracea"
/db_xref="taxon:3712"
/clone="g125a07"
/clone_lib="Bobuds01"
/note="Vector: M13 for .x reads, pluscript for .b and .g reads; Site: 1; EcoRV; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000D3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA provided by Pablo Rablinowicz (CSHL) and shotgun library prepared in McCombie lab."

BASE COUNT 64 a 78 c 71 g 96 t

ORIGIN

Alignment Scores:

Pred. No.: 1.5e+03 Length: 309
Score: 38.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 77.55% Indels: 0
DB: 17 Gaps: 0

US-09-823-649A-5 (1-11) x BH743626 (1-309)

QY 1 LeuserValArgLeugly**ProValysglu 11
|||||
Db 114 ATATCGATCGACCTTTCAGAACCCGTCAGAGAA 82

RESULT 14

LOCUS AM477975 338 bp mRNA linear EST 09-JUL-2000
DEFINITION 17923 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AM477975
VERSION AM477975.1 GI:7048081

KEYWORDS EST.
SOURCE pig.
ORGANISM *Sus scrofa*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 384)
AUTHORS Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
and Keele, J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith RPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -mismatch 12 options.
PCR Primers
FORWARD: AGAAGACAGCTATGACCAT
BACKWARD: GTTTCACCTACACGACG
Plate: 8 row: P column: 8
Seq primer: ATTAGCTGACCTATGAC.
Location/Qualifiers
1. 384
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="Pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 87 a 84 c 95 g 72 t
ORIGIN
Alignment Scores:
Pred. No.: 1.64e+03 Length: 384
Score: 38.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 77.55% Indels: 0
DB: 10 Gaps: 0
US-09-823-649A-5 (1-11) x AM477975 (1-338)
QY 1 LeuserValArgLeucly**Proval 9
|||||
Db 323 CTGTCAGTCAAGTGGGCTCTCTCTGTG 297
RESULT 15
AZ465796 384 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION IM0275N19R Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0275N19 R, DNA sequence.
ACCESSION AZ465796
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 384)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.,
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0275 Row: N Column: 19
Seq primer: CACACAGGAACAGCTATGACG
Class: plasmid ends
High quality sequence stop: 384.
FEATURES
Source
Location/Qualifiers
1. 384
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0275N19"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b [AF129072.1]), a copy number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 108 a 96 c 89 g 91 t
ORIGIN
Alignment Scores:
Pred. No.: 1.86e+03 Length: 384
Score: 38.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 77.95% Indels: 0
DB: 17 Gaps: 0
US-09-823-649A-5 (1-11) x AZ465796 (1-384)
QY 1 LeuserValArgLeucly**ProvallyGCU 11
|||||
Db 310 CTGGGTAGACTGGAGCCAGTGAAGGAC 342

Search completed: January 21, 2003, 12:24:22
Job time : 1172.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds

(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649a-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXXE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODEL=frame+pn.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09823649/rnat.21012003.093151.24626/ftp.query.fasta.1.1393
-DB=Published.Applications.NA-QFMT-fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsam62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-HNR.MIN=0 -ALIGN=15 -MODE=LOCAL -CURRENT=ptc -NORM=ext -HEAVYSIDE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649.ecgn.1.1.77 -rnat.21012003.093151.24626
-KCP=6 -ICP=3 -NO_XLPEXT -NO_MMAP -LARGEQUERY -NBG_SCORES=0 -WAIT -IONGLOS
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database:

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	50.0	257	10	US-09-878-574-15531	Sequence 15531, A
2	50.0	415	10	US-09-878-574-4032	Sequence 4032, Ap
3	50.0	462	10	US-09-864-761-95	Sequence 95, Appl
4	50.0	618	10	US-09-864-761-16934	Sequence 16934, A

5	50.0	618	10	US-09-864-761-18952	Sequence 18952, A
6	50.0	688	9	US-10-040-739-303	Sequence 303, App
7	50.0	989	10	US-09-864-761-2212	Sequence 2212, Ap
8	50.0	1556	9	US-09-981-353-64	Sequence 64, Appl
9	50.0	1842	9	US-09-981-876-113	Sequence 113, App
10	50.0	1903	9	US-09-981-876-55	Sequence 55, Appl
11	50.0	2037	10	US-09-822-830A-25	Sequence 25, Appl
12	50.0	2281	10	US-09-799-777-101	Sequence 101, App
13	50.0	2793	10	US-09-880-107-3457	Sequence 3457, Ap
14	50.0	3501	9	US-09-938-842A-13	Sequence 13, Appl
15	50.0	3531	9	US-09-992-598-32	Sequence 32, Appl
16	50.0	3531	9	US-09-989-293A-32	Sequence 32, Appl
17	50.0	3531	9	US-09-989-735-32	Sequence 32, Appl
18	50.0	3531	9	US-09-990-444-32	Sequence 32, Appl
19	50.0	3531	9	US-09-989-730-32	Sequence 32, Appl
20	50.0	3531	9	US-09-990-436-32	Sequence 32, Appl
21	50.0	3531	9	US-09-991-181-32	Sequence 32, Appl
22	50.0	3531	9	US-09-993-687-32	Sequence 32, Appl
23	50.0	3531	9	US-09-989-734-32	Sequence 32, Appl
24	50.0	3531	9	US-09-997-653-32	Sequence 32, Appl
25	50.0	3531	10	US-09-989-722-32	Sequence 32, Appl
26	50.0	3531	10	US-09-989-723-32	Sequence 32, Appl
27	50.0	3531	10	US-09-989-279-32	Sequence 32, Appl
28	50.0	3531	10	US-09-989-727-32	Sequence 32, Appl
29	50.0	3531	10	US-09-989-731-32	Sequence 32, Appl
30	50.0	3531	10	US-09-989-732-32	Sequence 32, Appl
31	50.0	3531	10	US-09-991-073-32	Sequence 32, Appl
32	50.0	3531	10	US-09-990-442-32	Sequence 32, Appl
33	50.0	3531	10	US-09-991-163-32	Sequence 32, Appl
34	50.0	3531	10	US-09-993-604-32	Sequence 32, Appl
35	50.0	3531	10	US-09-990-456-32	Sequence 32, Appl
36	50.0	3531	10	US-09-989-721-32	Sequence 32, Appl
37	50.0	4100	9	US-09-813-718-5	Sequence 5, Appl1
38	50.0	4273	9	US-09-974-298-143	Sequence 143, App
39	50.0	5174	9	US-09-813-718-1	Sequence 1, Appl1
40	50.0	13444	10	US-09-764-877-3660	Sequence 3660, Ap
41	50.0	17967	10	US-09-764-848-50	Sequence 50, Appl
42	50.0	18871	10	US-09-764-847-1317	Sequence 1317, Ap
43	44.4	82	10	US-09-878-574-7261	Sequence 7261, Ap
44	44.4	97	10	US-09-864-761-20170	Sequence 20170, A
45	44.4	194	9	US-09-900-379-107	Sequence 107, App

ALIGNMENTS

RESULT 1
US-09-878-574-15531
Sequence 15531, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: BYTUM, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 15531
LENGTH: 257
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701070216H1
US-09-878-574-15531

Alignment Scores:

Pred. No.:	2.91	Length:	257
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9

Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0
US-09-823-649A-1 (1-11) x US-09-878-574-15531 (1-257)
QY 1 Leu*****Glu 11
DB 8 CTGACCTCAGCGCGCGCGCGCACACAGAA 40
RESULT 2
US-09-878-574-4032
Sequence 4032, Application US/09878574
Patent No. US2002010546A1
GENERAL INFORMATION:
APPLICANT: Hyrum, Joseph R.
APPLICANT: Ia Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1998-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 4032
LENGTH: 415
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-002-Q1-B1-E10
US-09-878-574-4032
Alignment Scores:
Pred. No.: 3.59 Length: 415
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
Gaps: 0
US-09-823-649A-1 (1-11) x US-09-878-574-4032 (1-415)
QY 1 Leu*****Glu 11
DB 13 CTGACCTCAGCGCGCGCGCGCACACAGAA 45
RESULT 3
US-09-864-761-95/C
Sequence 95, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,667
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 95
LENGTH: 462
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007051.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN H474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
US-09-864-761-95
Alignment Scores:
Pred. No.: 3.77 Length: 462
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
Gaps: 0
US-09-823-649A-1 (1-11) x US-09-864 761-95 (1-462)
QY 1 Leu*****Glu 11
DB 288 CTGACGAGCAGCGCTTCACATCAGTAGTAA 256
RESULT 4
US-09-864-761-16934/C
Sequence 16934, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04

;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 16934
;; LENGTH: 618
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC007051.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN HONE MARROW, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN HONE MARROW, SIGNAL = 2.2
;; OTHER INFORMATION: NT HIT: AB037784.1, EVALUO 0.00e+00
;; OTHER INFORMATION: SWISSPROT HIT: P22760, EVALUO 2.00e-42
;; OTHER INFORMATION: EST_HUMAN HIT: BE154994.1, EVALUO 0.00e+00
US-09-864-761-16934
Alignment Scores:
Pred. No.: 4.28 Length: 618
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0
US-09-823-649a-1 (1-11) x US-09-864-761-16934 (1-618)
QY 1 Leu*****Glu 11
|||
DB 239 CTGACAGCAGCAGCCTTCCACATCAGTGA 207
RESULT 5
US-09-864-761-18952

;; Sequence 18952, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Acomica X-1
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 18952
;; LENGTH: 618
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC007919.18
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN HONE MARROW, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;; OTHER INFORMATION: NT HIT: AB037784.1, EVALUO 0.00e+00
;; OTHER INFORMATION: EST_HUMAN HIT: BE154994.1, EVALUO 0.00e+00
US-09-864-761-18952
Alignment Scores:
Pred. No.: 4.28 Length: 618

Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-1 (1-11) x US-09-864-761-18952 (1-618)

OY 1 Leu*****Glu 11
|||
DB 324 CTGCACGACGAGCCTCTCCATCATCAGTGA 356

RESULT 6
US-10-040-739-303/c

; Sequence 303, Application US/10040739

; Patent No. US20020173635A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John

LaValle, Edward

Racie, Lisa

Merberg, David

Treacy, Maurice

Spaulding, Vikki

TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS

NUMBER OF SEQUENCES: 1519

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

A PPLICATION NUMBER: US/10/040,739

FILING DATE: 07-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/036,520

FILING DATE: 03-JUN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 303:

SEQUENCE CHARACTERISTICS:

LENGTH: 683 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 303:

US-10-040-739-303

Alignment Scores:

Pred. No.: 4.47 Length: 683

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 9 Gaps: 0

US-09-823-649a-1 (1-11) x US-10-040-739-303 (1-683)

OY 1 Leu*****Glu 11
|||

DB 451 TTAGCTGCCACATCTCTCTCTCTCTCTCCGAG 419

RESULT 7

US-09-864-761-2212

; Sequence 2212, Application US/09864761

; Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecm1ca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing mgine vers. 1.1

SEQ ID NO 2212

LENGTH: 969

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC007919.18

OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.5

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3

US-09-864-761-2212

Alignment Scores:

```
Pred. No.: 5.27 Length: 589
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-1 (1-11) x US-09-864-761-2212 (1-989)
QY 1 Leu*****Glu 11
DB 676 CTGACAGCAGCAGCTCTTCACATCAAGTCAA 708

RESULT 8
US-09-981-353-64
; Sequence 64, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 1556
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 3231154CB1
US-09-981-353-64

Alignment Scores:
Pred. No.: 6.43 Length: 1556
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-1 (1-11) x US-09-981-353-64 (1-1556)
QY 1 Leu*****Glu 11
DB 1054 TTGGCATCATCATCTACTGCTGCTGCTGCCGAG 1086

RESULT 9
US-09-981-876-113/c
; Sequence 113, Application US/09981876
; Patent No. US2002016469A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/046,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
```


;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/056,886
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,877
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,889
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,893
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,630
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,878
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,662
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,872
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,882
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,637
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,903
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,888
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,879
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,880
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,894
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,911
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,636
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,874
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,910
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,864
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,631
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,845
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,892
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/047,595
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/057,761
;; PRIOR FILING DATE: 05-Sep-1997
;; PRIOR APPLICATION NUMBER: 60/047,599
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,588
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,585
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,586
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,590
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,594
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,589
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,593
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,614
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/043,578
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,576
;; PRIOR FILING DATE: 1997-04-11

;; PRIOR APPLICATION NUMBER: 60/047,501
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/043,670
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/056,632
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,664
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,876
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,881
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,909
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,875
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,862
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,887
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,908
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/048,964
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/057,650
;; PRIOR FILING DATE: 1997-09-05
;; PRIOR APPLICATION NUMBER: 60/056,884
;; PRIOR FILING DATE: 1997-08-22
;; NUMBER OF SEQ ID NOS: 280
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 113
;; LENGTH: 1842

Alignment Scores:
Bred. No.: 6.93 Length: 1842
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 9 Gaps: 0

US-09-823-649a-1 (1-11) x US-09-981-876-113 (1-1842)

QY 1 Leu*****gln 11
DB 164 CFACACAGCGCAGCAGCAGCAGCAGCA 132

RESULT 10
US-09-981-876-55/c
; Sequence 55, Application US/09/981876
; Patent No. US2002164669A1
; GENERAL INFORMATION:
; APPLICANT: Rosgen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001PI
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/043,670
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/056,632
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,664
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,876
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,881
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,909
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,875
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,862
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,887
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,908
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/048,964
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/057,650
PRIOR FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: 60/056,884
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 55
LENGTH: 1903

Alignment Scores:
Pred. No.: 7.03
Score: 9.00
Percent Similarity: 18.18%
Best Local Similarity: 18.18%
Query Match: 50.00%
DB: 9
Gaps: 0

US-09-823-649A-1 (1-11) x US-09-881-876-55 (1-1903)

Cy 1 Leu*****Glu 11
Db 203 CTACACACGCCGACGACGACGACGACGACAA 171

RESULT 11
US-09-822-830A-25
Sequence 25 Application US/09822830A
Patent No. US20020142952A1
GENERAL INFORMATION:
APPLICANT: Genetics Institute, Inc.
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6402
CURRENT APPLICATION NUMBER: US/09/822,830A
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,604
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 631
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 2037
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-830A-25

Alignment Scores:
Pred. No.: 7.24
Score: 9.00
Percent Similarity: 18.18%
Best Local Similarity: 18.18%
Query Match: 50.00%
DB: 10
Gaps: 0

US-09-823-649A-1 (1-11) x US-09-822-830A-25 (1-2037)

Cy 1 Leu*****Glu 11
Db 95 TTGGCCGCCGCCGCCGCCGCCGCCGCCGACG 127

RESULT 12
US-09-799-777-101/c
Sequence 101 Application US/09799777
Patent No. US2002091244A1
GENERAL INFORMATION:
APPLICANT: Ial, Proeti
Billman, Jennifer L.
Corley, Neil C.
Guegler, Karl J.
Baugh, Mariah
Sather, Susan
Shah, Purvi
TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
NUMBER OF SEQUENCES: 154
CORRESPONDENCE ADDRESS:
ADDRESS: JINGYI PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/799,777
FILING DATE: 06-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/002,485
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LARRY J.
REGISTRATION NUMBER: 36,749
REFERENCE/WORKSET NUMBER: 0459 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
FAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 2281 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRA1W0712
CLONE: 1615809
SEQUENCE DESCRIPTION: SEQ ID NO: 101 :
US-09-799-777-101

Alignment Scores:
Pred. No.: 7.61
Score: 9.00
Percent Similarity: 18.18%
Best Local Similarity: 18.18%
Query Match: 50.00%
DB: 10
Gaps: 0

US-09-823-649A-1 (1-11) x US-09-799-777-101 (1-2281)

QY 1 Leu*****Glu 11

DB 695 TTAGTGCACATCTTCTCTCTCCGAG 663

RESULT 13

US-09-880-107-3437

; Sequence 3437, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darci T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3437

; LENGTH: 2793

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U79725

US-09-880-107-3437

Alignment Scores:

Pred. No.: 8.32 Length: 2793

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 10 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-880-107-3437 (1-2793)

QY 1 Leu*****Glu 11

DB 1099 TTGGCATCATCTACTGCTCTCTCCGAG 1131

RESULT 14

US-09-938-842A-13/C

; Sequence 13, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kieps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,667

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 13

; LENGTH: 3501

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-13

Alignment Scores:

Pred. No.: 9.19 Length: 3501

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 9 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-938-842A-13 (1-3501)

QY 1 Leu*****Glu 11

DB 332 CTCATCCTCTCTCTACATCTTCACAA 300

RESULT 15

US-09-992-598-32

; Sequence 32, Application US/09992598

; Patent No. US20020160384A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Bolstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerlitsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kiljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Thomas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2730P1020

; CURRENT APPLICATION NUMBER: US/09/992,598

; CURRENT FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

1	PRIOR FILING DATE: 1998-06-02	1	PRIOR APPLICATION NUMBER: 60/089907
2	PRIOR APPLICATION NUMBER: 60/087759	2	PRIOR FILING DATE: 1998-06-18
3	PRIOR FILING DATE: 1998-06-02	3	PRIOR APPLICATION NUMBER: 60/089947
4	PRIOR APPLICATION NUMBER: 60/087827	4	PRIOR FILING DATE: 1998-06-19
5	PRIOR FILING DATE: 1998-06-03	5	PRIOR APPLICATION NUMBER: 60/089948
6	PRIOR APPLICATION NUMBER: 60/088021	6	PRIOR FILING DATE: 1998-06-19
7	PRIOR FILING DATE: 1998-06-04	7	PRIOR APPLICATION NUMBER: 60/089952
8	PRIOR APPLICATION NUMBER: 60/088025	8	PRIOR FILING DATE: 1998-06-19
9	PRIOR FILING DATE: 1998-06-04	9	PRIOR APPLICATION NUMBER: 60/090246
10	PRIOR APPLICATION NUMBER: 60/088026	10	PRIOR FILING DATE: 1998-06-22
11	PRIOR FILING DATE: 1998-06-04	11	PRIOR APPLICATION NUMBER: 60/090252
12	PRIOR APPLICATION NUMBER: 60/088028	12	PRIOR FILING DATE: 1998-06-22
13	PRIOR FILING DATE: 1998-06-04	13	PRIOR APPLICATION NUMBER: 60/090254
14	PRIOR APPLICATION NUMBER: 60/088029	14	PRIOR FILING DATE: 1998-06-22
15	PRIOR FILING DATE: 1998-06-04	15	PRIOR APPLICATION NUMBER: 60/090349
16	PRIOR APPLICATION NUMBER: 60/088030	16	PRIOR FILING DATE: 1998-06-23
17	PRIOR FILING DATE: 1998-06-04	17	PRIOR APPLICATION NUMBER: 60/090355
18	PRIOR APPLICATION NUMBER: 60/088033	18	PRIOR FILING DATE: 1998-06-23
19	PRIOR FILING DATE: 1998-06-04	19	PRIOR APPLICATION NUMBER: 60/090429
20	PRIOR APPLICATION NUMBER: 60/088326	20	PRIOR FILING DATE: 1998-06-24
21	PRIOR FILING DATE: 1998-06-04	21	PRIOR APPLICATION NUMBER: 60/090431
22	PRIOR APPLICATION NUMBER: 60/088367	22	PRIOR FILING DATE: 1998-06-24
23	PRIOR FILING DATE: 1998-06-05	23	PRIOR APPLICATION NUMBER: 60/090435
24	PRIOR APPLICATION NUMBER: 60/088402	24	PRIOR FILING DATE: 1998-06-24
25	PRIOR FILING DATE: 1998-06-05	25	PRIOR APPLICATION NUMBER: 60/090444
26	PRIOR APPLICATION NUMBER: 60/088212	26	PRIOR FILING DATE: 1998-06-24
27	PRIOR FILING DATE: 1998-06-05	27	PRIOR APPLICATION NUMBER: 60/090445
28	PRIOR APPLICATION NUMBER: 60/088217	28	PRIOR FILING DATE: 1998-06-24
29	PRIOR FILING DATE: 1998-06-05	29	PRIOR APPLICATION NUMBER: 60/090472
30	PRIOR APPLICATION NUMBER: 60/088655	30	PRIOR FILING DATE: 1998-06-24
31	PRIOR FILING DATE: 1998-06-09	31	PRIOR APPLICATION NUMBER: 60/090535
32	PRIOR APPLICATION NUMBER: 60/08734	32	PRIOR FILING DATE: 1998-06-24
33	PRIOR FILING DATE: 1998-06-10	33	PRIOR APPLICATION NUMBER: 60/090540
34	PRIOR APPLICATION NUMBER: 60/08738	34	PRIOR FILING DATE: 1998-06-24
35	PRIOR FILING DATE: 1998-06-10	35	PRIOR APPLICATION NUMBER: 60/090542
36	PRIOR APPLICATION NUMBER: 60/08742	36	PRIOR FILING DATE: 1998-06-24
37	PRIOR FILING DATE: 1998-06-10	37	PRIOR APPLICATION NUMBER: 60/090557
38	PRIOR APPLICATION NUMBER: 60/08810	38	PRIOR FILING DATE: 1998-06-24
39	PRIOR FILING DATE: 1998-06-10	39	PRIOR APPLICATION NUMBER: 60/090676
40	PRIOR APPLICATION NUMBER: 60/08824	40	PRIOR FILING DATE: 1998-06-25
41	PRIOR FILING DATE: 1998-06-10	41	PRIOR APPLICATION NUMBER: 60/090678
42	PRIOR APPLICATION NUMBER: 60/08826	42	PRIOR FILING DATE: 1998-06-25
43	PRIOR FILING DATE: 1998-06-10	43	PRIOR APPLICATION NUMBER: 60/090690
44	PRIOR APPLICATION NUMBER: 60/08858	44	PRIOR FILING DATE: 1998-06-25
45	PRIOR FILING DATE: 1998-06-11	45	PRIOR APPLICATION NUMBER: 60/090694
46	PRIOR APPLICATION NUMBER: 60/08861	46	PRIOR FILING DATE: 1998-06-25
47	PRIOR FILING DATE: 1998-06-11	47	PRIOR APPLICATION NUMBER: 60/090695
48	PRIOR APPLICATION NUMBER: 60/08876	48	PRIOR FILING DATE: 1998-06-25
49	PRIOR FILING DATE: 1998-06-11	49	PRIOR APPLICATION NUMBER: 60/090696
50	PRIOR APPLICATION NUMBER: 60/089105	50	PRIOR FILING DATE: 1998-06-25
51	PRIOR FILING DATE: 1998-06-12	51	PRIOR APPLICATION NUMBER: 60/090862
52	PRIOR APPLICATION NUMBER: 60/089440	52	PRIOR FILING DATE: 1998-06-26
53	PRIOR FILING DATE: 1998-06-16	53	PRIOR APPLICATION NUMBER: 60/090863
54	PRIOR APPLICATION NUMBER: 60/089512	54	PRIOR FILING DATE: 1998-06-26
55	PRIOR FILING DATE: 1998-06-16	55	PRIOR APPLICATION NUMBER: 60/091360
56	PRIOR APPLICATION NUMBER: 60/089514	56	PRIOR FILING DATE: 1998-07-01
57	PRIOR FILING DATE: 1998-06-16	57	PRIOR APPLICATION NUMBER: 60/091478
58	PRIOR APPLICATION NUMBER: 60/089532	58	PRIOR FILING DATE: 1998-07-02
59	PRIOR FILING DATE: 1998-06-17	59	PRIOR APPLICATION NUMBER: 60/091544
60	PRIOR APPLICATION NUMBER: 60/089538	60	PRIOR FILING DATE: 1998-07-01
61	PRIOR FILING DATE: 1998-06-17	61	PRIOR APPLICATION NUMBER: 60/091519
62	PRIOR APPLICATION NUMBER: 60/089598	62	PRIOR FILING DATE: 1998-07-02
63	PRIOR FILING DATE: 1998-06-17	63	PRIOR APPLICATION NUMBER: 60/091626
64	PRIOR APPLICATION NUMBER: 60/089599	64	PRIOR FILING DATE: 1998-07-02
65	PRIOR FILING DATE: 1998-06-17	65	PRIOR APPLICATION NUMBER: 60/091633
66	PRIOR APPLICATION NUMBER: 60/089600	66	PRIOR FILING DATE: 1998-07-02
67	PRIOR FILING DATE: 1998-06-17	67	PRIOR APPLICATION NUMBER: 60/091978
68	PRIOR APPLICATION NUMBER: 60/089653	68	PRIOR FILING DATE: 1998-07-07
69	PRIOR FILING DATE: 1998-06-17	69	

Alignment Scores:

Pred. No.:	9.22	Length:	3531
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	9	Gaps:	0

US-09-823-649a-1 (1-11) x US-09-992-598-32 (1-3531)

QY 1 Leu*****Glu 11

|||

Db 74 CTTCCAGCTGCAGCAGCTGCAGCAGCTGCAGAG 106

Search completed: January 21, 2003, 10:07:39
Job time : 31.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 : Search time 1168.14 seconds

(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-1

Perfect score: 18

Sequence: 1 LXXXXXXX 11

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters: -v -
-MODEL=frame+ p2n.model -DEV=slp
-O=/cgn2_1/USPTO.spool/US09823649.r/unat_21012003_093151_24616/app-query.fasta.1.1393
-DB=slp -OPMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.ccl -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-DUPM=ptc -NORM=ext -HFM=ext -HFM=ext -HFM=ext -HFM=ext -HFM=ext -HFM=ext
-USER=US09823649.eccn_1.1.2820.etrnat_21012003_093151_24616 -ICPU=3
-NO_XLPHY -NO_MMP -LARGEDQUERY -NGS_SCORES=0 -WAIT -LONGLOG -DHY_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlum:*
5: em_estlov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hlc:*
9: qb_est1:*
10: qb_est2:*
11: qb_hlc:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estlom:*
17: qb_gss:*
18: em_gss_hum:*
19: em_gss_lin:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	50.0	110	13	BM289398
2	9	50.0	112	12	BM642522
3	9	50.0	112	12	BE842524
4	9	50.0	113	10	AM389253
5	9	50.0	113	12	BE842537
6	9	50.0	114	12	BE842532
7	9	50.0	114	12	BE842534
8	9	50.0	127	17	AQ427805
9	9	50.0	132	10	AM389255
10	9	50.0	147	14	B0636167
11	9	50.0	151	14	D28324
12	9	50.0	157	12	BE842527
13	9	50.0	170	9	AA383829
14	9	50.0	181	12	BE842526
15	9	50.0	193	12	BE842523
16	9	50.0	203	12	B1903789
17	9	50.0	230	12	BF803350
18	9	50.0	237	12	HP883878
19	9	50.0	240	14	C46922
20	9	50.0	247	12	BE6467560
21	9	50.0	249	9	AU076866
22	9	50.0	250	12	BE6468040
23	9	50.0	254	12	BE6467571
24	9	50.0	256	14	B0777127
25	9	50.0	259	14	BQ978708
26	9	50.0	267	10	AV869931
27	9	50.0	270	12	CNS07FEB
28	9	50.0	273	13	BM18560
29	9	50.0	278	9	AL134269
30	9	50.0	281	14	B0162056
31	9	50.0	287	10	BM567353
32	9	50.0	294	10	BE261459
33	9	50.0	297	10	AV981681
34	9	50.0	300	9	AJ461594
35	9	50.0	300	9	AU099432
36	9	50.0	300	14	C11627
37	9	50.0	300	14	C31277
38	9	50.0	300	14	C51118
39	9	50.0	301	14	C55175
40	9	50.0	301	10	AM610337
41	9	50.0	304	9	A1370697
42	9	50.0	310	10	AV947349
43	9	50.0	321	14	BM289398
44	9	50.0	322	14	BM755663
45	9	50.0	324	13	BM407825

ALIGNMENTS

RESULT 1
BM289398
LOCUS
DEFINITION
Sequence 5 Oryza sativa Nootripathu Differential Display Oryza
sativa cDNA, mRNA sequence.
ACCESSION
BM289398
VERSION
BM289398.1
GI:21245160
KEYWORDS
EST.
ORGANISM
Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 110)
RajA.J.A.J., Ravendran,M. and Sadasiyam,S.

TITLE Differentially Expressed Sequence Tags in the roots of drought tolerant rice variety Nootripathu during water stress

JOURNAL Unpublished (2001)

COMMENT Contact: Raja, J. A. J.; Raveendran, M.; Sadasivam, S.
Centre for Plant Molecular Biology
Tamil Nadu Agricultural University
Coimbatore - 641003, India
Email: jaja@ayahoo.com

FEATURES
source location/Qualifiers
1. .110
/organism="Oryza sativa"
/cultivar="Nootripathu"
/db_xref="taxon:4530"
/clone_lib="Oryza sativa Nootripathu Differential Display"
/tissue_type="Root tissue"
/dev_stage="Vegetative"
/note="Differentially expressing mRNAs were identified by differential display of total RNA isolated from the roots of normal and water stressed (70% leaf relative water content) plants. The upregulation of the ESTs during drought was confirmed by northern hybridization analyses."

BASE COUNT 27 a 27 c 20 g 35 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	2.18e-05	Length:	110
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	13	Gaps:	0

US-09-823-649a-1 (1-11) x BM289398 (1-110)

Oy 1 leu*****Glu 11
|||

DB 8 CTTACCGCACTACAGCGCGCTACGCGAG 40

RESULT 2
BE842522/c 112 bp mRNA linear EST 22-SEP-2000
LOCUS IL3-ST0142-070600-167-A10 ST0142 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE842522
ACCESSION BE842522.1 GI:10274900
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 112)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

REFERENCE
AUTHORS
CONTACT: Simpson A.J.G.
LABORATORY OF Cancer Genetics
LUDWIG Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ptl=st2-IL3-ST0142-070600-167-A10&f3=2000-06-07&f4=1)
Seq primer: puc 18 forward

JOURNAL
MEDLINE
COMMENT

FEATURES
source location/Qualifiers
1. .112
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0142"
/dev_stage="Adult"
/note="Organ: stomach; Vector: puc18; Site:1: Sma1; Site:2: Sma1; A mini-library was made by cloning products derived from OKSITES PCR (U.S. letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 13 a 45 c 34 g 20 t

ORIGIN

Alignment Scores:

Pred. No.:	2.22e-05	Length:	112
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649a-1 (1-11) x BE842522 (1-112)

Oy 1 leu*****Glu 11
|||

DB 34 TTACGACAGCAGCAGCAGCAGCAGCAGCAG 2

RESULT 3
BE842524 112 bp mRNA linear EST 22-SEP-2000
LOCUS IL3-ST0142-070600-167-C04 ST0142 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE842524
ACCESSION BE842524.1 GI:10274902
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 112)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

REFERENCE
AUTHORS
CONTACT: Simpson A.J.G.
LABORATORY OF Cancer Genetics
LUDWIG Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ptl=st2-IL3-ST0142-070600-167-C04&f3=2000-06-07&f4=1)
Seq primer: puc 18 forward
High quality sequence stop: 112.

JOURNAL
MEDLINE
COMMENT

FEATURES
source location/Qualifiers
1. .112
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0142"
/dev_stage="Adult"

QY 1 Leu*****Glu 11
 DB 35 TTACGACGACGACGACGACGATCGAG 3

RESULT 6
 BE842532/c 114 bp mRNA linear EST 22-SEP-2000
 LOCUS
 DEFINITION I13-ST0142-070600-168-F05 ST0142 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE842532
 VERSION BE842534.1 GI:10274910
 KEYWORDS EST.

ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 114)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 /M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

AUTHORS
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPSP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=612-l13-ST0142-070
 600-168-F05&l3=2000-06-07&l4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 114.

FEATURES
 Location/Qualifiers
 source 1..114
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="ST0142"
 /dev_stage="Adult"
 /note="Organ: stomach; Vector: puc18; Site: 1: Smal;
 Site: 2: Smal; A mini-library was made by cloning products
 derived from ORESTES PCR (O.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 14 a 46 c 34 g 20 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,26e-05 Length: 114
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 12 Gaps: 0

US-09-823-649a-1 (1-11) x BE842532 (1-114)

QY 1 Leu*****Glu 11
 DB 36 TTACGACGACGACGACGATCGAG 4

RESULT 7
 BE842534/c 114 bp mRNA linear EST 22-SEP-2000
 LOCUS

DEFINITION I13-ST0142-070600-168-G05 ST0142 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE842534
 VERSION BE842534.1 GI:10274912
 KEYWORDS EST.

ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 114)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 /M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

AUTHORS
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPSP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=612-l13-ST0142-070
 600-168-G05&l3=2000-06-07&l4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 114.

FEATURES
 Location/Qualifiers
 source 1..114
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="ST0142"
 /dev_stage="Adult"
 /note="Organ: stomach; Vector: puc18; Site: 1: Smal;
 Site: 2: Smal; A mini-library was made by cloning products
 derived from ORESTES PCR (O.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 14 a 46 c 34 g 20 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,26e-05 Length: 114
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 12 Gaps: 0

US-09-823-649a-1 (1-11) x BE842534 (1-114)

QY 1 Leu*****Glu 11
 DB 36 TTACGACGACGACGACGATCGAG 4

RESULT 8
 A0427805/c 127 bp DNA linear GSS 24-MAR-1999
 LOCUS A0427805
 DEFINITION CITR1-El Homo sapiens genomic clone 2562E1, DNA
 sequence.
 ACCESSION A0427805
 VERSION A0427805.1 GI:4500712
 KEYWORDS GSS.
 ORGANISM human.
 Homo sapiens

Pred. No.:	2.98e-05	Length:	151
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	14	Gaps:	0

3

PREFACE

AA383829/c 170 bp mRNA linear EST 21-APR-1997
 LOCUS AA383829
 DEFINITION EST97265 Testis I Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA383829
 VERSION AA383829.1 GI:2036147
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 170)
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulton,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-ali,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Geoghegan,N.S., Glick,A., Gnehm,C.L., Hanna,M.C., Hedblom,B., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Geo,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wel,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
 TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavet@igf.org
 For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hg1/hg1.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 FEATURES
 source 1..170
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):188010"
 /db_xref="taxon:9606"
 /clone_lib="Testis I"
 /sex="male"
 /dev_stage="adult"
 /note="Organ: testis; Vector: pBluescript SK-; Site_1: EcoRI, Site_2: XhoI"
 BASE COUNT 36 a 71 c 32 g 29 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.36e-05 Length: 170
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-823-649a-1 (1-11) x AA383829 (1-170)
 QY 1 Leu*****Glu 11
 DB 135 CTTAGCGCTGCTGCGCGAGCTGCCGAA 103
 RESULT 14
 BE842526
 LOCUS BE842526 181 bp mRNA linear EST 22-SEP-2000

IL3-ST0142-070600-167-F04 ST0142 Homo sapiens cDNA, mRNA sequence.
 BE842526
 LOCUS BE842526 181 bp mRNA linear EST 22-SEP-2000
 DEFINITION IL3-ST0142-070600-167-B06 ST0142 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE842523
 VERSION BE842523.1 GI:10274901
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 181)
 AUTHORS Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Britous,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jorgensen,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?IL3-ST0142-070600-167-F04&f3-2000-06-07&f4-1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 181.
 Location/Qualifiers
 FEATURES
 source 1..181
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="ST0142"
 /dev_stage="Adult"
 /note="Organ: stomach; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (O.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 29 a 58 c 65 g 29 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.57e-05 Length: 181
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 12 Gaps: 0
 US-09-823-649a-1 (1-11) x BE842526 (1-181)
 QY 1 Leu*****Glu 11
 DB 79 TTGACGACGACGACGACGCGATCGGAG 111
 RESULT 15
 BE842523
 LOCUS BE842523 193 bp mRNA linear EST 22-SEP-2000
 DEFINITION IL3-ST0142-070600-167-B06 ST0142 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE842523
 VERSION BE842523.1 GI:10274901
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 193)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Br<11
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL:
(http://www.ludwig.org.br/scripts/gethtml2.pl?r1=6t2=103-S70142-070
600-167-B066t3=2000-06-07&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 193.
Location/Qualifiers

FEATURES

source

1. 193
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="S70142"
/dev_stage="Adult"
/notes="Organ: stomach; Vector: puc18; Site_1: Sma1;
Site_2: Sma1; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 26 a 76 c 61 g 30 t
ORIGIN

Alignment Scores:

Pred. No.:	3.81e-05	Length:	193
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-1 (1-11) x BE842523 (1-193)

QY 1 leu*****Ciu 11
|||
DB 79 TTAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 111

Search completed: January 21, 2003, 12:24:10
Job time : 1171.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 21, 2003, 09:31:57 : Search time 1168.14 seconds
(without alignments)
152.507 Million cell updates/sec

Title: us-09-823-649a-7
Perfect score: 49
Sequence: 1 LAQNINIKRKE 11

Scoring table:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 33208132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -p2n model -DRV-xlp
-MODBL=frame+ p2n model /usr982369/umat.21012003.093151.24616/app.query.fasta_1.1393
-Q/cgn2.1/USPFO.spool/US982369/umat.21012003.093151.24616/app.query.fasta_1.1393
-DB-EST -QFMT=fastlap -SUFTX-1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX-DIsum2 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=1 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Plto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US982369 @CGL 1.1.2820 @umat.21012003.093151.24616 -NCPV=6 -ICPV=3
-NO_XLPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -MAIT -LONGLOG -DRV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	85.7	695	10	BB468375
2	40	81.6	359	13	BB468375
3	39	81.6	501	13	B1492651
4	39	79.6	832	17	AA0855611
5	38	77.6	449	12	BF635786
6	38	77.6	1034	17	AC155700
7	37	75.5	125	12	BF371247
8	37	75.5	197	13	BM026044
9	37	75.5	261	9	AV245515
10	37	75.5	427	14	B0612936
11	37	75.5	520	17	AA0732941
12	37	75.5	528	17	AA0708467
13	37	75.5	675	17	AC178233
14	37	75.5	718	17	AC172746
15	37	75.5	738	10	AV709998
16	37	75.5	753	12	BE821761
17	37	75.5	756	10	AW350371
18	37	75.5	761	17	AA0745123
19	37	75.5	819	17	BB248796
20	37	75.5	838	14	B0441565
21	37	75.5	888	17	BH158636
22	37	75.5	1084	17	BH133865
23	37	75.5	1155	17	AG039702
24	36	73.5	271	13	B1681170
25	36	73.5	378	17	AA391764
26	36	73.5	401	17	AA754534
27	36	73.5	431	17	BH754529
28	36	73.5	458	9	AA084503
29	36	73.5	554	17	AA072741
30	36	73.5	618	13	BH070175
31	36	73.5	620	14	B0603915
32	36	73.5	629	10	AA616886
33	36	73.5	665	17	AA0316963
34	36	73.5	680	10	BB639624
35	36	73.5	745	17	BB639624
36	36	73.5	745	17	AA0316963
37	36	73.5	745	17	AA0316963
38	36	73.5	745	17	AA0316963
39	36	73.5	745	17	AA0316963
40	36	73.5	745	17	AA0316963
41	36	73.5	745	17	AA0316963
42	36	73.5	745	17	AA0316963
43	36	73.5	745	17	AA0316963
44	36	73.5	745	17	AA0316963
45	36	73.5	745	17	AA0316963

ALIGNMENTS

RESULT 1
BB468375/C
LOCUS BB468375
DEFINITION BB468375 RIKEN full-length enriched, 12 days embryo eyecall Mus
ACCESSION BB468375
VERSION BB468375
KEYWORDS BB468375.2 GI:16427508
SOURCE
ORGANISM Mus musculus
EST
house mouse.
Mus musculus
Bukariyotai; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 695)
REFERENCE
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanaigaki,T., Igarashi,A.,
Hiramoto,K., Hoti,F., Ishii,Y., Ito,M., Kawada,J., Kono,H., Kouda

Query Match: 81.638 Indels: 0
DB: 13 Gaps: 0

US-09-823-649A-7 (1-11) x B1495153 (1-359)

QY 1 LeuAlagInasInLeuAsnIle***Arglys 10
|||||
DB 266 TTGGCCCAAAACCTTAATAAAGGAAA 295

RESULT 3
B1492651

LOCUS B1492651 501 bp mRNA linear EST 28-AUG-2001

DEFINITION B1492651.1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:2484498 3', mRNA sequence.

ACCESSION B1492651

VERSION B1492651.1 GI:153131995

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 501)
Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R. and Morton,C.C.
Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening
Genomics 23, 42-50 (1994)

TITLE Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and Reproductive Biology
Brigham and Women's Hospital
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
Tel: 617 732 7980
Fax: 617 738 6996
Email: cmorton@rics.bwh.harvard.edu
DNA sequencing and analyses were performed by National Institutes of Health Intramural Sequencing Center (NISC) see <http://www.nisc.nih.gov>.
This clone is available royalty-free through LAMN; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Plate: LAM6182 row: F column: 19
Seq primer: T7 primer.

FEATURES
source Location/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2484498"
/clone_lib="Morton Fetal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ear; Vector: pBluescript SK-; 42-50 (1994) Cloned unidirectionally. Primer: Oligo dT. Fetal cochlea, normal. 3/8 of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR Vector. Library constructed by N. Robertson, C. Morton.
-5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'

BASE COUNT 145 a 110 c 81 g 165 t

ORIGIN

Alignment Scores:
Pred. No.: 73 Length: 501
Score: 40.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 81.638 Indels: 0
DB: 13 Gaps: 0

US-09-823-649A-7 (1-11) x B1492651 (1-501)

QY 1 LeuAlagInasInLeuAsnIle***Arglys 10

DB 263 TTGGCCCAAAACCTTAATAAAGGAAA 292
|||||

RESULT 4
A0855611

LOCUS A0855611 832 bp DNA linear GSS 12-MAY-2000

DEFINITION CpG1705A Cp10WAgDNA1 Cryptosporidium parvum genomic DNA sequence.

ACCESSION A0855611

VERSION A0855611.1 GI:6206068

KEYWORDS GSS.

SOURCE Cryptosporidium parvum.

ORGANISM Cryptosporidium parvum
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Cryptosporidiidae; Cryptosporidium.

REFERENCE 1 (bases 1 to 832)
Strong W.B. and Nelson R.G.
Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)

TITLE Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry
San Francisco General Hospital-University of California, San Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8846
Fax: 415 206 3353
Email: malaria@itsa.ucsf.edu
For Annotation data see <http://medsfrh.ucsf.edu/ld/CpTags/home.html>
Seq primer: M13(-21) forward
Class: shotgun.

FEATURES
source Location/Qualifiers
1..832
/organism="Cryptosporidium parvum"
/strain="10WA"
/db_xref="taxon:5607"
/clone_lib="Cp10WAgDNA1"
/lab_host="E. coli XL2 Blue MRF"
/note="Vector: pBluescript II (SK-); site: 1: EcoRV, C. parvum (10WA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorsen of the Stanford DNA Sequencing and Technology Center (<http://sequence-www.stanford.edu/group/leachdev/sear.htm>). The randomly sheared gDNA was chromatographed on Sepharyl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoRV-digested, alkaline phosphatase-treated pBluescript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF-. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

BASE COUNT 335 a 114 c 196 g 185 t 2 others

ORIGIN

Alignment Scores:
Pred. No.: 250 Length: 832
Score: 39.00 Matches: 7
Percent Similarity: 90.00% Conservative: 2
Best Local Similarity: 70.00% Mismatches: 1
Query Match: 79.59% Indels: 0
DB: 17 Gaps: 0

US-09-823-649A-7 (1-11) x A0855611 (1-832)

QY 2 AlagInasInLeuAsnIle***ArglysGlu 11
|||||

DB 626 GCTCGAAGCTGAACTTACGAAAAAGAA 655

RESULT 5
BF635786

LOCUS BF635786 449 bp mRNA linear EST 19-DEC-2000

DEFINITION NF039609DT11069 Drought Medicago truncatula cDNA clone NF039609DT

5', mRNA sequence.
 ACCESSION BP635786
 VERSION BP635786.1 GI:11899944
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 449)
 REFERENCE
 AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula drought library
 JOURNAL Unpublished (2000)
 COMMENT
 CONTACT: May GD
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7391
 Fax: 580 221 7380
 Email: gdmay@noble.org
 Insert Length: 449 Std Error: 0.00
 Plate: 039 row: G column: 09
 Seq primer: TCACACAGCAACACGCTATGAC.
 FEATURES
 source
 1..449
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF039509PT"
 /clone_lib="Drought"
 /tissue_type="Plantlets"
 /dev_stage="Pooled timepoints"
 /note="Vector: lambda Zap. Contains a mixture of entire
 plantlets harvested in a series of days-post-watering
 timepoints."
 BASE COUNT 126 a 74 c 118 g 131 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 172 Length: 449
 Score: 38.00 Matches: 8
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 1
 Query Match: 77.55% Indels: 0
 DB: 12 Gaps: 0
 US-09-823-649a-7 (1-11) x BP635786 (1-449)
 QY 1 LeuAAGlnAsnLeuAsnIle**Arg 9
 |||||
 Db 347 TTGGCCCAAAATTGACATCAGAG 373
 RESULT 6
 LOCUS AG155700 1034 bp DNA linear GSS 09-JAN-2002
 DEFINITION Pan troglodytes DNA, clone: RP43-020B14.TJ, genomic survey
 sequence.
 ACCESSION AG155700
 VERSION AG155700.1 GI:16685378
 KEYWORDS GSS.
 SOURCE Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
 Male BAC Library clone:RP43-020B14.TJ.
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 REFERENCE
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of Library RPCI-43
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1034)

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suenho-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimpesegsc.riken.go.jp, URL:http://npg.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
 end was generated during the Rad process and may have higher chance
 of clone tracking errors.
 PRIMERS
 Sequencing: TJ
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI
 location/Qualifiers
 1..1034
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="RP43-020B14.TJ"
 /sex="male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-43 Chimpanzee Male BAC library"
 BASE COUNT 306 a 202 c 282 g 235 t 9 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 567 Length: 1034
 Score: 38.00 Matches: 8
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 1
 Query Match: 77.55% Indels: 0
 DB: 17 Gaps: 0
 US-09-823-649a-7 (1-11) x AG155700 (1-1034)
 QY 2 AAGlnAsnLeuAsnIle**ArgLys 10
 |||||
 Db 468 GCACAAATTATACATTAAACGAAA 494
 RESULT 7
 LOCUS BP371247 175 bp mRNA linear EST 24-NOV-2000
 DEFINITION KC6-FN0138-260700-011-A01 FN0138 Homo sapiens CDNA, mRNA sequence.
 ACCESSION BP371247
 VERSION BP371247.1 GI:1133376
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 AUTHORS 1 (bases 1 to 175)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunschweiler, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL
 MEDLINE
 COMMENT
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-27049922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC6&l2=RC6-FN0138-260700-011-A01&t3=2000-07-26&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 175.

FEATURES

source

1. 175
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0138"
/dev_stage="Adult"
/note="Organ: prostate, normal; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESSES PCR (U.S. letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 38 a 48 c 28 g 61 t

ORIGIN

Alignment Scores:

Pred. No.: 74.7 Length: 175
Score: 37.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 75.51% Indels: 0
DB: 12 Gaps: 0

US-09-823-649a-7 (1-11) x BF371247 (1-175)

OY 1 LeuAlaGlnAsnIleuAsnIle***ArglyscIu 11

DB 50 CTAGTCACACCTTAATATCTCCACAGACAT 82

RESULT 8

BM026044 187 bp mRNA linear EST 30-OCT-2001
LOCUS fu77e06.x1 Gong zebrafish testis Dantio rerio cDNA clone 5376875 3'
DEFINITION similar to SM:FP36.HUMAN 064252 EUKARYOTIC TRANSLATION INITIATION
FEATURE 3 SUBUNIT 6 ; , mRNA sequence.

ACCESSION

EM026044
F0026044.1 GI:16540400

VERSION

1.0

KEYWORDS

zebrafish.

SOURCE

Dantio rerio

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 187)

AUTHORS

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

, K., Steptoe, M., Reising, B., Allen, M., Bowers, Y., Peterson, B.,

Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schalk, R., Ritter, E.,

Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.

and Wilson, R.

Washu zebrafish EST Project 1998

Unpublished (1998)

contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1800

Email: zbrafish@wustl.edu

The library was constructed by Dr. Z. Gong. DNA sequencing by:

Washington University Genome Sequencing Center St. Louis. Please

contact Zhiyuan Gong for further information on this library

(National University of Singapore, Department of Biological

Sciences, Lower Kent Ridge Road, Singapore 119260).

Trace considered overall poor quality

High quality sequence stop: 1.

Location/Qualifiers

source

1. 187
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="5376875"
/clone_lib="Gong zebrafish testis"
/sex="male"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis (pooled); Vector: pBluescript SK-; Site.1: XhoI; Site.2: EcoRI; Poly A+ RNA was isolated from the testes of 31 male adult zebrafish (4-5 month old). cDNAs were made using oligo-dT primers and inserted into lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo mass-excision to pBluescript SK- following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). please contact Zhiyuan Gong for further information on this library (National University of Singapore, Department of Biological Sciences, Lower Kent Ridge Road, Singapore 119260)."

BASE COUNT 51 a 46 c 42 g 48 t

ORIGIN

Alignment Scores:
Pred. No.: 82.1 Length: 187
Score: 37.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 75.51% Indels: 0
DB: 13 Gaps: 0

US-09-823-649a-7 (1-11) x BM026044 (1-187)

OY 1 LeuAlaGlnAsnIleuAsnIle***ArglyscIu 11

DB 70 ATGGCGATACCTTAATATCTCTGGAGAG 38

RESULT 9

AV245515 261 bp mRNA linear EST 04-NOV-1999
LOCUS AV245515 RIKEN full-length enriched, 0 day neonate head Mus
DEFINITION musculus cDNA clone 48314019 3' similar to D50263 Mus musculus
DAN mRNA, mRNA sequence.

ACCESSION

AV245515
AV245515.1 GI:6232974

VERSION

1.0

KEYWORDS

zebrafish.

SOURCE

house mouse

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 261)

REFERENCE

1 (bases 1 to 261)

AUTHORS

Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,

Ishii, Y., Ishikawa, T., Itoh, M., Izawa, K., Kadota, K., Kagawa, T., Kai

, C., Kawali, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,

Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,

Owa, C., Ozawa, Y., Saito, H., Sano, K., Sato, K., Shibata, K., Shibata

, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suahara, Y., Suzuki, H.,

Suzuki, H., Takahashi, F., Tateno, M., Tomioka, N., Tsunoda, Y.,

Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T.,

Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al. 1999)

Unpublished (1999)

contact: Yoshinori Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp

URL: <http://genome.res.riken.go.jp/>

Sasaki, N., Izawa, M., Watanabe, M., Tanaka, K., Tanaka, T., Yoneda, Y.,

FEATURES	Location/Qualifiers
source	1. .261

BASE COUNT	51 a	72 c	51 g	87 u
ORIGIN				

US-09-823-649A-7 (1-11) x AV245515 (1-261)

```

QY      1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
          |||||  |||||  |||||  |||||
Db      70 CTAGCCCAAGGTTAAATTACTAGGAAGAA 38

```

RESULT 10			
B0612936			
LOCUS			
DEFINITION	B0612936	427 bp	linear mRNA
DEFINITION	sap8h07.y1	Gm-c1087	Glycine max cDNA clone SOYB8N CLONE 10:
DEFINITION	Gm-c1087-6901	5'	similar to SW:DAPF_SVN3 P74657 DIAMINOPYMELEATE
DEFINITION	EPIMERASE ;	mRNA sequence.	
DEFINITION			EST 26-JUN-2002

ACCESSION B0612936
VERSION B0612936.1
KEYWORDS GI:21602605
SOURCE EST.
ORGANISM soybean.
CLICING max
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Eumariophyta: Magnoliophyta: eudicotyledons: core eudicots:

REFERENCE
AUTHORS
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 427)
Shoemaker, R., Keim, P., Vockin, L., Erpelidg, J., Coryell, V., Khanna

TITLE	JOURNAL	COMMENT
Public Soybean EST Project	Unpublished (1999)	Contact: Shoemaker R/Public Soybean EST Project

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-537-4363 or contact: cut@resgen.com web site:
www.resgen.com
putative full length read
vector to vector length is 471
Seq primer: -40RP from G10C0.

FEATURES
source

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SORBEN CLONE ID: Gm-c1087-6901"
/clone_lib="Gm-c1087"
/issue_type="Roots"
/lab_host="DH10B"

```

Note: "vector: phagescript II SK(-); site1: EcoRI; site2: XhoI). The mRNA was prepared using poly(A)tract mRNA system from PROMEGA. The cDNA was prepared using the STRATAGENE kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dt) sequence with a XhoI restriction site (5'-CAGACAGACGACGACGACGACACTAGCTCCAGCTTTTTCCTTTTTTTTTT-3'). PfuRI adapters (5'-OH-AATTTCGCCACGAG and 3'-CGCCGTCGCP) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the PfuRI XhoI restriction sites of the phagescript vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH5B by electroporation. Library is in JM medium with ampicillin and glycerol 8%. Average insert size: 800 bp."

BASE COUNT	109 a	78 c	113 g	127 l
ORIGIN				

Pred. No.:	267	Length:	42
Score:	37.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	75.51%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649A-7 (1-11) x BQ612936 (1-427)

```

QY      1 LeuAlaGlnAsnLeuAsnIle***Arg 9
          |||||
Db      63 TTGGCCCAAAATTGAACTCATGAGC 8
          |||||
RESULT 11
AQ732941

```

LOCUS	A0732941	520 bp	DNA	linear	GenSeq 15-JUL-1992
DEFINITION	HS-554c_B1_H01_17A RCR1-11 Human Male BAC Library Homo sapiens genomic clone Plate=1124 Col=1 Row=J, DNA sequence.				
ACCESSION	A0732941				
VERSION	A0732941.1	GI:5504493			
KEYWORDS	GSS.				

US-09-823-649A-7 (1-11) x AV709998 (1-738)

OY 1 LeuAlaGlnAsnLeuAsnIle**ArgIysGlu 11
||| |||||
Db 367 CTTGCCAGAAATTTGACATTTATCATGAGGGA 399

Search completed: January 21, 2003, 12:24:30
Job time : 1172.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds

(without alignments)
165,854 Million cell updates/sec

Title: US-09-823-649A-7
Sequence: 1 LAONLNIKKE 11

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters: "-
-MODEL=frame+g2n.model -DEV=xlp
-O/cg2n2_1/USPto.spool/US0982369/unal.21012003.093151.24626/app_query.fasta.1.1393
-DB=Published.Applications.NA -OPM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -SCAP=1 -BND=1 -MATRIX=ilos62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -MOR=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09823649 @CGN.1.1_77 @unal.21012003.093151.24626
-NCPu=6 -ICPU=3 -NO.XLPHY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Published.Applications.NA:*

1: /cg2n2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2n2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2n2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2n2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2n2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2n2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2n2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2n2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2n2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cg2n2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2n2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cg2n2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cg2n2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cg2n2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	39	79.6	2545	10	US-09-974-300-1524
2	37	75.5	3177	10	US-09-801-574-3
3	36	73.5	181	10	US-09-960-352-13109
4	36	73.5	2649	10	US-09-815-242-6571

5	36	73.5	2787	10	US-09-815-242-6320	Sequence 6320, Ap
6	36	73.5	8033	10	US-09-070-927A-121	Sequence 121, App
7	35	71.4	151	10	US-09-878-574-8417	Sequence 8417, Ap
8	35	71.4	492	10	US-09-783-590-6288	Sequence 6288, Ap
9	35	71.4	780	10	US-09-910-943-561	Sequence 561, App
10	34	69.4	239	10	US-09-867-701-8965	Sequence 8965, Ap
11	34	69.4	684	10	US-09-974-300-6024	Sequence 6024, Ap
12	34	69.4	735	10	US-09-815-242-8265	Sequence 8265, Ap
13	34	69.4	753	10	US-09-815-242-4354	Sequence 4354, Ap
14	34	69.4	1664	10	US-09-939-980-259	Sequence 259, App
15	34	69.4	2670	10	US-09-815-242-9138	Sequence 9138, App
16	34	69.4	4754	10	US-09-982-091A-1	Sequence 1, App1
17	33	67.3	399	10	US-09-878-574-2036	Sequence 2036, Ap
18	33	67.3	439	10	US-09-892-206-6	Sequence 6, App1
19	33	67.3	597	10	US-09-864-761-12301	Sequence 12301, A
20	33	67.3	1458	9	US-10-047-412A-3	Sequence 3, App1
21	33	67.3	1458	9	US-09-880-107-2424	Sequence 2424, Ap
22	33	67.3	1536	9	US-09-938-842A-222	Sequence 222, App
23	33	67.3	2170	9	US-10-047-412A-14	Sequence 14, App1
24	33	67.3	2361	10	US-09-880-107-2271	Sequence 2271, Ap
25	33	67.3	2416	10	US-09-748-739A-3	Sequence 3, App1
26	33	67.3	2416	10	US-09-748-739A-5	Sequence 5, App1
27	33	67.3	2416	10	US-09-748-739A-7	Sequence 7, App1
28	33	67.3	2416	10	US-09-748-739A-16	Sequence 16, App1
29	33	67.3	2657	10	US-09-892-206-1	Sequence 1, App1
30	33	67.3	5000	12	US-10-063-763-4	Sequence 4, App1
31	32	65.3	446	10	US-09-867-701-7987	Sequence 7987, App
32	32	65.3	695	10	US-09-764-877-716	Sequence 716, App
33	32	65.3	768	9	US-09-938-842A-1461	Sequence 1461, Ap
34	32	65.3	995	12	US-10-021-509-4	Sequence 4, App1
35	32	65.3	995	12	US-10-021-509-12	Sequence 12, App1
36	32	65.3	1278	10	US-09-770-445-22	Sequence 22, App1
37	32	65.3	1730	9	US-10-028-072-7	Sequence 7, App1
38	32	65.3	1802	10	US-09-908-711-132	Sequence 132, App
39	32	65.3	1966	10	US-09-887-776-40	Sequence 40, App1
40	32	65.3	2000	9	US-09-938-842A-4751	Sequence 4751, Ap
41	32	65.3	2845	10	US-09-070-927A-462	Sequence 462, App
42	32	65.3	3600	10	US-09-764-877-3729	Sequence 3729, Ap
43	32	65.3	3600	10	US-09-764-877-3730	Sequence 3730, Ap
44	32	65.3	8194	9	US-09-799-629-15	Sequence 15, App1
45	32	65.3	176373	9	US-10-095-407-17	Sequence 17, App1

ALIGNMENTS

RESULT 1
US-09-974-300-1524
; Sequence 1524, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974.300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680.598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279.526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1524
; LENGTH: 2545
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1524
Alignment Scores:
Pred. No.: 14.6
Score: 39.00
Percent Similarity: 81.82%
Length: 2545
Matches: 8
Conservative: 1

Best Local Similarity: 72.73% Mismatches: 2
Query Match: 79.59% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-974-300-1524 (1-2545)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
||||:||||||| ||| |||||
Db 2170 TTGTCCCAAACTGGGATTTACAGAAAGGAA 2202

RESULT 2
US-09-801-574-3
; Sequence 3, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijing Jeremy
; APPLICANT: Page, David C.
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-801-574-3

Alignment Scores:
Pred. No.: 53.9 Length: 3177
Score: 37.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 75.51% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-801-574-3 (1-3177)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
||||:||||||| ||| |||||
Db 1880 ATAAGCCAGATCTGATATATAAAACAAAGAA 1912

RESULT 3
US-09-960-352-13109
; Sequence 13109, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengding
; APPLICANT: Byatt, John C.
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13109
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 56-LIB3057-023-Q1-K1-P8
US-09-960-352-13109

Alignment Scores:
Pred. No.: 2.89 Length: 181
Score: 36.00 Matches: 7
Percent Similarity: 90.00% Conservative: 2

Best Local Similarity: 70.00% Mismatches: 1
Query Match: 73.47% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-960-352-13109 (1-181)

QY 2 AlaGlnAsnLeuAsnIle**ArgLysGlu 11
||||:||||||| ||| |||||
Db 72 GCAAAAGATCTGATATATACAGACAGCA 101

RESULT 4
US-09-815-242-6571
; Sequence 6571, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Onisen, Karl H.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6571
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2649)
US-09-815-242-6571

Alignment Scores:
Pred. No.: 72.9 Length: 2649
Score: 36.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 73.47% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-815-242-6571 (1-2649)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
||||:||||||| ||| |||||
Db 2179 TTGTCCCAAACTTAGGCACTTACTAGAAAGCA 2211

RESULT 5
US-09-815-242-6320
; Sequence 6320, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert

```

? APPLICANT: Ohlsen, Karl L.
? APPLICANT: Zyskind, Judith W.
? APPLICANT: Wall, Daniel
? APPLICANT: Trawick, John D.
? APPLICANT: Carr, Grant J.
? APPLICANT: Yamamoto, Robert T.
? APPLICANT: Xu, H. Howard
? TITLE OF INVENTION: Identification of Essential Genes in
? FILE OF INVENTION: Prokaryotes
? FILE REFERENCE: ELITRA.011A
? CURRENT APPLICATION NUMBER: US/09/815,242
? CURRENT FILING DATE: 2001-03-21
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? NUMBER OF SEQ ID NOS: 14110
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 6320
? LENGTH: 2787
? TYPE: DNA
? ORGANISM: Escherichia coli
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(2787)
US-09-815-242-6320

Alignment Scores:
Pred. No.: 77.5 Length: 2787
Score: 36.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 73.47% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-815-242-6320 (1-2787)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArqLysGlu 11
Db 2317 CTGGCCGCGCATTCACATTCACGTAAGAA 2349

RESULT 6
US-09-070-927A-121
? Sequence 121, Application US/09070927A
? Patent No. US20020120116A1
? GENERAL INFORMATION:
? APPLICANT: Charles A. Kunsch
? Patrick J. Dillon
? Steven Barash
? TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
? NUMBER OF SEQUENCES: 982
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Human Genome Sciences, Inc.
? STREET: 9410 Key West Avenue
? CITY: Rockville
? STATE: Maryland
? COUNTRY: USA
? ZIP: 20850
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
? COMPUTER: HP Vectra 486/33
? OPERATING SYSTEM: MSDOS version 6.2
? SOFTWARE: ASCII Text
```

```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/070,927A
? FILING DATE: 04-May-2000
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/046,555
? FILING DATE: 1997-09-16
? APPLICATION NUMBER: 60/044,031
? FILING DATE: 1997-05-06
? APPLICATION NUMBER: 60/066,009
? FILING DATE: 1997-11-14
? ATTORNEY/AGENT INFORMATION:
? NAME: Kenley K. Hoover
? REGISTRATION NUMBER: 40,302
? REFERENCE/DOCKET NUMBER: PB369
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (301) 309-8504
? TELEFAX: (301) 309-8512
? INFORMATION FOR SEQ ID NO: 121:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 8033 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 121:
US-09-070-927A-121

Alignment Scores:
Pred. No.: 277 Length: 8033
Score: 36.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 0
Query Match: 73.47% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-070-927A-121 (1-8033)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArqLysGlu 11
Db 7018 TTCTCTCAAAACCTACCATTCATCAGCAAGCAA 7050

RESULT 7
US-09-878-574-8417
? Sequence 8417, Application US/09878574
? Patent No. US20020110548A1
? GENERAL INFORMATION:
? APPLICANT: Byrum, Joseph R.
? APPLICANT: La Rose, Thomas J.
? APPLICANT: Thompson, Michael D.
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
? FILE REFERENCE: 38-21(15401)A
? CURRENT APPLICATION NUMBER: US/09/878,574
? CURRENT FILING DATE: 2001-12-21
? PRIOR APPLICATION NUMBER: 09/353,535
? PRIOR FILING DATE: 1999-06-14
? NUMBER OF SEQ ID NOS: 15775
? SEQ ID NO 8417
? LENGTH: 151
? TYPE: DNA
? ORGANISM: Glycine max
? OTHER INFORMATION: Clone ID: 701101161H1
US-09-878-574-8417

Alignment Scores:
Pred. No.: 3.91 Length: 151
Score: 35.00 Matches: 7
Percent Similarity: 80.00% Conservative: 1
Best Local Similarity: 70.00% Mismatches: 2
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-878-574-8417 (1-151)
```

```
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgIys 10
|||||
Db 93 TTGAGCAGAACTGAACTTGCACCGAAGA 122

RESULT 8
US-09-783-590-6288/c
; Sequence 6488, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2c1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6288
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (101)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (378)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (440)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (451)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (489)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-6288

Alignment Scores:
Pred. No.: 16.2 Length: 492
Score: 35.00 Matches: 6
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 1
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-783-590-6288 (1-492)

OY 3 GlnAsnLeuAsnIle***ArgIysGlu 11
|||||
Db 297 CAAATATGAATCTTATTCGGAAGAAG 271

RESULT 9
US-09-910-943-561
; Sequence 561, Application US/09910943

; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/1G1480S1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 561
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(780)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-561

Alignment Scores:
Pred. No.: 28.2 Length: 780
Score: 35.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 71.43% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-910-943-561 (1-780)

OY 1 LeuAlaGlnAsnLeuAsnIle***ArgIysGlu 11
|||||
Db 132 TTAGCGAGAAATTAACTATCATCTGTAACAA 164

RESULT 10
US-09-867-701-8966/c
; Sequence 8966, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aiale, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Hartlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: PatentIn Version 4.0
; SEQ ID NO 8966
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-8966

Alignment Scores:
Pred. No.: 11.4 Length: 239
Score: 34.00 Matches: 7
Percent Similarity: 72.73% Conservative: 1
Best Local Similarity: 63.64% Mismatches: 3
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-867-701-8966 (1-239)

OY 1 LeuAlaGlnAsnLeuAsnIle***ArgIysGlu 11
|||||
Db 78 CTTGCACGAGAACTTGCCTTTCATAGAAAGAA 46

RESULT 11
US-09-974-300-6024
; Sequence 6024, Application US/09974300
; Patent No. US20020146721A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Betka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085,500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: F: USEQ for Windows Version 4.0
; SEQ ID NO 6024
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6024

Alignment Scores:
Pred. No.: 40.5 Length: 684
Score: 34.00 Matches: 7
Percent Similarity: 81.82% Conservative: 2
Best Local Similarity: 63.64% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-974-300-6024 (1-684)
Qy 1 LeuAlagInasInleuAsnile**Arglysglu 11
|||||:|||||:|||||:|||||:
Db 218 CTCGCTCAAAAGCCTACGCTTACGAGAAAGAA 250

RESULT 12
US-09-815-242-8266/c
; Sequence 8266, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8266
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1)...(735)
US-09-815-242-8266

Alignment Scores:
Pred. No.: 44.1 Length: 735
Score: 34.00 Matches: 6
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 54.55% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-815-242-8266 (1-735)
Qy 1 LeuAlagInasInleuAsnile**Arglysglu 11
|||||:|||||:|||||:|||||:
Db 307 CTCGCTCAAAATCTAAATTTAAATATATCGTCAT 275

RESULT 13
US-09-815-242-4354/c
; Sequence 4354, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4354
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4354

Alignment Scores:
Pred. No.: 45.4 Length: 735
Score: 34.00 Matches: 6
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 54.55% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-815-242-4354 (1-735)
Qy 1 LeuAlagInasInleuAsnile**Arglysglu 11
|||||:|||||:|||||:|||||:
Db 340 CTCGCTCAAAATCTAAATTTAAATATATCGTCAT 308

RESULT 14
```

US-09-939-980-259
; Sequence 259, Application US/09939980
; Patent No. US2002008234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; Burnham, Martin
; Hodgson, John
; Knowles, David
; Lonetto, Michael
; Nicholas, Richard
; Pratt, Julie
; Reichard, Richard
; Rosenberg, Martin
; Ward, Judith
; TITLE OF INVENTION: No. US2002008234A1el Prokaryotic Polynucleotides,
; Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimml, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 259:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 259:
US-09-939-980-259
Alignment Scores:
Pred. No.: 118 Length: 1664
Score: 34.00 Matches: 6
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 54.55% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0
US-09-823-649a-7 (1-11) x US-09-939-980-259 (1-1664)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
|||||
Db 329 CTCGCTCAAAATCTTAATTAAATAATCGTAT 361
RESULT 15
US-09-815-242-9138
; Sequence 9138, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes
; FILE REFERENCE: EILVRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9138
; LENGTH: 2670
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2670)
US-09-815-242-9138
Alignment Scores:
Pred. No.: 208 Length: 2670
Score: 34.00 Matches: 7
Percent Similarity: 72.73% Conservative: 1
Best Local Similarity: 63.64% Mismatches: 3
Query Match: 69.39% Indels: 0
DB: 10 Gaps: 0
US-09-823-649a-7 (1-11) x US-09-815-242-9138 (1-2670)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
|||||
Db 2200 TTCTCTAATAATCTTCCGATTTACGCTAACGAA 2232

Search completed: January 21, 2003, 10:08:29
Job time : 31.5714 secs

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/54, 9/12, C12Q 1/68	A1	(11) International Publication Number: WO 98/40496 (43) International Publication Date: 17 September 1998 (17.09.98)
(21) International Application Number: PCT/US98/05095 (22) International Filing Date: 12 March 1998 (12.03.98) (30) Priority Data: 60/039,610 12 March 1997 (12.03.97) US (71) Applicant: THE PERKIN-ELMER CORPORATION [US/US]; 850 Lincoln Centre Drive, Foster City, CA 94404 (US). (72) Inventors: BRANDIS, John; 106 Sheffield, Hercules, CA 94547 (US). BLOOM, Curtis; 2631 Chalet Place, Chino Hills, CA 91709 (US). RICHARDS, Jack; 677 Deodar Lane, Bradbury, CA 91010 (US). (74) Agent: BORTNER, Scott, D.; The Perkin-Elmer Corporation, 850 Lincoln Centre Drive, Foster City, CA 94404 (US).		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: DNA POLYMERASES HAVING IMPROVED LABELED NUCLEOTIDE INCORPORATION PROPERTIES (57) Abstract The present invention relates to mutant DNA polymerases that exhibit reduced discrimination against labeled nucleotides into polynucleotides. The DNA polymerases of the invention have at least one mutation in the nucleotide label interaction region of the enzyme such the mutation results in reduced discrimination against labeled nucleotides. The nucleotide label interaction regions is located at portions of the O-helix, (ii) the K-helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. In addition to providing novel mutant DNA polymerases, the invention also provides polynucleotides encoding the subject mutant DNA polymerases. The polynucleotides provided may comprise expression vectors for the recombinant production of the mutant polymerases. The invention also provides host cells containing the subject polynucleotides. The invention also includes numerous methods of using the subject DNA polymerases, including uses for chain termination sequencing and PCR. Another aspect of the invention is to provide kits for synthesizing fluorescently labeled polynucleotides in accordance with the methods of the invention. Kits of the invention comprise a mutant DNA polymerase of the invention and a fluorescently labeled nucleotide that exhibits reduced discrimination with respect to the mutant DNA polymerase in the kit. <div style="text-align: right;">E681</div>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**DNA POLYMERASES HAVING IMPROVED LABELED
NUCLEOTIDE INCORPORATION PROPERTIES**

5 Inventors: John Brandis, Curtis Bloom, and Jack Richards

Field of the Invention

10 The invention is related to DNA polymerases having mutations that alter the ability of the enzyme to incorporate labeled nucleotides into a polynucleotide molecule.

Background

15 DNA polymerases are enzymes that synthesize the formation of DNA molecules from deoxynucleotide triphosphates using a template DNA strand and a complementary synthesis primer annealed to a portion of the template. A detailed description of DNA polymerases and their enzymological characterization can be found in Kornberg, DNA Replication Second Edition, W. H. Freeman (1989).

20 DNA polymerases have a variety of uses in molecular biology techniques suitable for both research and clinical applications. Foremost among these techniques are DNA sequencing and nucleic acid amplification techniques such as PCR (polymerase chain reaction).

25 The amino acid sequence of many DNA polymerases have been determined. Sequence comparisons between different DNA polymerase have identified many regions of homology between the different enzymes. X-ray diffraction studies have determined the tertiary structures of Klenow fragment, T7 DNA polymerase, and Taq DNA polymerase. Studies of the tertiary structures of DNA polymerases and amino acid sequence comparisons have revealed numerous structural similarities between diverse DNA polymerases. In general, DNA polymerases have a large cleft that is thought to accommodate the binding of duplex DNA. This cleft is formed by two sets of helices, the first set is referred to as the "fingers" region and the second set of helices is referred to as the "thumb" region. The bottom of the cleft is formed by anti-parallel β sheets and is referred to as the "palm" region. Reviews of DNA polymerase structure can be found in Joyce and Steitz, Ann. Rev. Biochem. 63:777-822 (1994). Computer readable data files describing the three-dimensional structure of some

30

DNA polymerases have been publicly disseminated.

Fluorescently labeled nucleotides have greatly simplified and improved the utility of many procedures in molecular biology. The use of fluorescently labeled nucleotides for labeling polynucleotides in synthesis procedures, has to a large extent replaced the use of radioactive labeling. Fluorescently labeled nucleotides have been widely used in DNA sequencing, see Smith et al Nature 321:674-679 (1986), in PCR, and other forms of polynucleotide fragment analysis.

A major problem with using fluorescently labeled nucleotides is the ability of DNA polymerases to discriminate against the incorporation of fluorescently labeled nucleotides. For example, the inventors have discovered that in competition assays between a TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein) labeled 2' 3' dideoxynucleotide and the corresponding unlabeled dideoxynucleotide, Taq DNA polymerase incorporates the unlabeled dideoxynucleotide into DNA at least 85 times more frequently than the corresponding unlabeled nucleotide. This discrimination between labeled and unlabeled nucleotides has profound effects on procedures using DNA polymerases to label DNA. For example, much larger amounts of fluorescently labeled nucleotide must be used in sequencing reactions. This large amount of fluorescently labeled nucleotide is expensive and can generate excessive background fluorescence, thereby reducing the yield of sequence information.

In view of the problems arising from the ability of DNA polymerases to discriminate against the incorporation of fluorescently labeled nucleotides, the inventors have developed several novel DNA polymerases that have reduced discrimination against the incorporation of one or more fluorescently labeled nucleotides into DNA.

Summary

Naturally occurring DNA polymerases preferentially incorporate unlabeled nucleotides over corresponding fluorescently labeled nucleotides into polynucleotides. This ability of DNA polymerases to discriminate against fluorescently labeled nucleotide has undesirable effects on many molecular biology procedures that require the enzymatic addition of fluorescently labeled nucleotides, e.g., labeled dideoxy terminator sequencing. The present invention relates to mutant DNA polymerases that exhibit reduced discrimination against fluorescently labeled nucleotides into polynucleotides.

The DNA polymerases of the invention have at least one mutation in the nucleotide

label interaction region of the enzyme such that the mutation results in reduced discrimination against fluorescently labeled nucleotides. The nucleotide label interaction region of a DNA polymerase is formed by portions of the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. Amino acid residues within the nucleotide label interaction region as defined by TET (II) •ddC are E520, A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832. The sites at R660, T664, and E681 are of preferred sites for introducing mutations. In a preferred embodiment of the invention for use with fluorescein-type dyes, a mutation is present at position 681 converting an E (glutamic acid) to M (methionine), i.e., E681M. In a preferred embodiment of the invention for use with fluorescein- fluorescein energy transfer dyes a mutation is present at position 657 converting an L (leucine) to a G (glycine). In addition to providing mutant Taq DNA polymerases having reduced discrimination against labeled nucleotides, the invention includes mutants derived from a wide variety of DNA polymerases, both thermostable and otherwise.

In addition to providing novel mutant DNA polymerases, the invention also provides polynucleotides encoding the subject mutant DNA polymerases. The polynucleotides provided may comprise expression vectors for the recombinant production of the mutant polymerases. The invention also includes host cells containing the subject polymerase polynucleotides.

The invention also includes numerous methods of using the subject DNA polymerases. The subject methods involve synthesizing a fluorescently labeled polynucleotide by means of a polynucleotide synthesis reaction catalyzed by a mutant DNA polymerase that has reduced discrimination against incorporating labeled nucleotides into polynucleotides. The subject methods of polynucleotide synthesis include the step of extending a primed polynucleotide template with at least one fluorescent labeled nucleotide, wherein the extension is catalyzed by a DNA polymerase that has reduced discrimination against labeled nucleotides into polynucleotides. The subject methods of synthesizing a fluorescently labeled polynucleotide may be used in a variety of methods such as Sanger sequencing and the polymerase chain reaction (PCR).

Another aspect of the invention is to provide kits for synthesizing fluorescently labeled polynucleotides in accordance with the methods of the invention. Kits of the invention comprise a mutant DNA polymerase of the invention and a fluorescently labeled nucleotide that exhibits reduced discrimination with respect to the mutant DNA polymerase in the kit.

5

Brief Description of the Drawings

Figure 1 is a computer model of DNA bound to Taq DNA polymerase. Amino acid residues that form the nucleotide label interaction site are highlighted in orange. The rest of the polymerase is indicated in green. The template is indicated in blue. The dye moiety of the labeled nucleotide is red. The remainder of the labeled nucleotide is white.

10

Figure 2 is plot of a next nucleotide effect assay.

Figure 3 is plot of a next nucleotide effect assay.

Figure 4 is a representation of the structure of the fluorescently labeled nucleotide "TET(II)•ddCTP."

15

Detailed Description of Specific Embodiments of the Invention.

Terminology

Positions of amino acid residues within a DNA polymerase are indicated by either numbers or number/letter combinations. The numbering starts at the amino terminus residue. The letter is the single letter amino acid code for the amino acid residue at the indicated position in the naturally occurring enzyme from which the mutant is derived. Unless specifically indicated otherwise, an amino acid residue position designation should be construed as referring to the analogous position in all DNA polymerases, even though the single letter amino acid code specifically relates to the amino acid residue at the indicated position in Taq DNA polymerase.

25

Individual substitution mutations are indicated by the form of a letter/number/letter combination. The letters are the single letter code for amino acid residues. The numbers indicate the amino acid residue position of the mutation site. The numbering system starts at the amino terminus residue. The numbering of the residues in Taq DNA polymerase is as described in U.S. Patent No. 5,079,352 (Gelfand). Amino acid sequence homology between different DNA polymerases permits corresponding positions to be assigned to amino acid residues for DNA polymerases other than Taq. Unless indicated otherwise, a given number

30

refers to position in Taq DNA polymerase. The first letter, i.e., the letter to the left of the number, represents the amino acid residue at the indicated position in the non-mutant enzyme. The second letter represents the amino acid residue at the same position in the mutant enzyme. For example, the term "R660D" indicates that the arginine at position 660 has been replaced by an aspartic acid residue.

The term "discrimination" as used herein refers to the property of a DNA polymerase to preferentially incorporate unlabeled nucleotides over corresponding fluorescently labeled nucleotides into DNA, i.e., the DNA polymerase discriminates against the fluorescently labeled nucleotide. Preferential incorporation may be measured in an assay in which a fluorescently labeled 2'3' dideoxynucleotide and a corresponding unlabeled 2'3' dideoxynucleotide compete for incorporation into the same site of a polynucleotide. An example of such an assay can be found below in example 2.

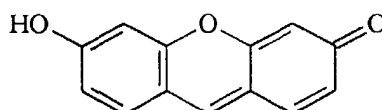
The term "reduced discrimination" as used herein refers to reduction in discrimination against incorporation of a fluorescently labeled nucleotides in a mutant DNA polymerase as compared to the parent enzyme. A reduction in discrimination may be described quantitatively by reference to the selectivity assays in Example 2 or reference to other assays providing for measurement of the same properties of the polymerase. A reduction in selectivity number as measured by the selectivity assays is a reduction in discrimination and may be expressed by a ratio of selectivity numbers. For example, a mutant DNA polymerase with a selectivity number of 8 would have a 10-fold reduction in discrimination when compared with a parent DNA polymerase having a selectivity number of 80.

The term "parent" or "parent enzyme" is used to distinguish a mutant DNA polymerase from the DNA polymerase that the mutant enzyme was derived from. Thus any naturally occurring DNA polymerase may be referred to as parent enzyme. A first DNA polymerase having mutations with respect to a naturally occurring enzyme is also be referred to as a parent enzyme with respect to a second DNA polymerase having additional mutations.

The term "discrimination reducing mutations" refers to mutations in the nucleotide label interaction region of a DNA polymerase that result in reduced discrimination against the incorporation of fluorescently labeled nucleotides. The term is used to distinguish mutations in a DNA polymerase, including mutations in the nucleotide label interaction region, that do not reduce discrimination against fluorescently labeled nucleotides from mutations that do reduce discrimination.

The term "nucleotide" as used herein, unless specifically noted otherwise, is used broadly to refer to both naturally occurring nucleotide and a variety of analogs including 2',3'-dideoxynucleotides.

The term "fluorescein-type dyes" refers to a class of xanthene dye molecules which include the following fused three-ring system:



where a wide variety of substitutions are possible at each deoxy ring position. A particularly preferred subset of fluorescein-type dyes include the 4,7,-dichlorofluoresceins (Menchen). Examples of fluorescein-type dyes used as fluorescent labels in DNA sequencing methods include 6-carboxyfluorescein (6-FAM), 5-carboxyfluorescein (5-FAM), 6 -carboxy-4,7,2',7'-tetrachlorofluorescein (TET), 6-carboxy-4,7,2',4',5',7'-hexachlorofluorescein (HEX), 5-(and 6)carboxy-4',5'-dichloro-2'7'-dimethoxyfluorescein (JOE), and 5-carboxy-2',4',5',7'-tetrachlorofluorescein (ZOE). Many times the designation -1 or -2 is placed after an abbreviation of a particular dye, e.g., HEX-1. The "-1" and "-2" (or "I" and "II") designations indicate the particular dye isomer being used. The 1 and 2 isomers are defined by the elution order (the 1 isomer being the first to elute) of free dye in a reverse-phase chromatographic separation system utilizing a C-8 column and an elution gradient of 15% acetonitrile/85% 0.1 M triethylammonium acetate to 35% acetonitrile / 65% 0.1 M triethylammonium acetate.

The term "alkynylamino type linker" refers to an alkynylamino linker of the type as described in U.S. Patent No. 5,047,519 (Hobbs), U.S. Patent No. 5,151,507 (Hobbs), and U.S. Patent Application No. 08/696,808, filed August 13, 1996. Additional alkynylamino type linkers are described in U. S. Patent Application No. 08/833,855, filed April 10, 1997.

The term "TET(II)-ddCTP" refers to the fluorescently labeled nucleotide of the structure indicated in figure 4.

The term "fluorescence energy transfer dye" refers to dye moieties joined by a linker that permits fluorescence energy transfer between the two dye moieties. For use in chain termination sequencing, the linker is sufficiently small and of the proper shape and orientation

to permit a DNA polymerase to incorporate a nucleotide triphosphate labeled with the dye interest. Examples of energy transfer dyes can be found in European Patent Application No. EP 0 805 140, U.S. Patent Application No. 08/642,330 (filed May 3, 1996), and U.S. Patent Application No. 08/726,462 (filed October 4, 1996).

5 The term "mutation" as used herein refers to a change in amino acid residue at a specific location of a protein. The change in amino acid residue is a change defined with respect to a naturally occurring protein. A protein having a mutation may be referred to as a "mutant" protein.

10 Embodiments of the Invention

 The present invention relates to DNA polymerases containing mutations that reduce the ability of the polymerase to discriminate against the incorporation of fluorescently labeled nucleotides into polynucleotides. These mutations are in a region of the DNA polymerase molecule referred to herein as "the nucleotide label interaction region." The nucleotide label
15 interaction region is formed by portions of three regions of the DNA polymerase. These three regions are located in (i) the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. DNA polymerases having reduced discrimination against fluorescently labeled nucleotides are particularly useful for chain termination DNA sequencing using 2'3' dideoxynucleotides, i.e., Sanger type
20 sequencing.

 Enzyme kinetic experiments (described in examples 2 and 3) performed with Taq DNA polymerase and fluorescently labeled dideoxynucleotides support a theory that Taq DNA polymerase and other DNA polymerases, undergo a conformational shift upon the binding of nucleotides during DNA synthesis. This predicted conformational shift suggests
25 a set of amino acid residues that interact with fluorescent labels joined by a linker to the nucleic acid base of a nucleotide, thereby resulting in discrimination against nucleotides that are fluorescently labeled. This set of amino acid residues forms the nucleotide label interaction region. The specific molecular model for the binding of fluorescently labeled nucleotide to a DNA polymerase proposed by the applicants is used to predict the amino acid
30 residues that form the nucleotide label interaction region of a given DNA polymerase. Applicants model for a conformational shift in DNA polymerase during DNA synthesis is offered as a explanation of how the nucleotide label interaction region was determined. The

model provides guidance in making mutations in DNA polymerase that reduce the ability of a DNA polymerase to discriminate against the incorporation of fluorescently labeled nucleotides into polynucleotides. Figure 1 is a computer model showing how DNA and Taq DNA polymerase interact in the model. Whether or not the true mechanism of DNA polymerase-nucleotide interaction is the same or different as the model used to determine the parameters of the nucleotide label interaction region is not determinative to the operability of the invention described herein.

The mutant DNA polymerases of the invention exhibit reduced discrimination against nucleotides labeled with a fluorescein-type dye. In other words, the mutant DNA polymerases of the invention contain at least one mutation that increases the ability of the polymerase to incorporate a fluorescein-type dye labeled nucleotide relative to the corresponding unlabeled nucleotide. In addition to reduced discrimination against nucleotides labeled with fluorescein-type dyes, the mutant DNA polymerases of the invention may also exhibit reduced discrimination against nucleotides labeled with other fluorescent dyes that are not fluorescein-type dyes, as well as reduced discrimination against other detectable moieties. The fluorescently labeled nucleotides for which a given embodiment of the mutant DNA polymerases of the invention exhibit reduced discrimination may vary with respect to the particular fluorescent label, the linker used to attach the fluorescent label to the nucleotide, the site of attachment for the linker on the fluorescent label, the specific nucleotide base that is selected, and the site of attachment for the linker on the nucleotide. The precise degree of reduction in discrimination against a fluorescently labeled nucleotide will vary in accordance with the specific mutation or mutations introduced into the DNA polymerase. The precise degree of reduction in discrimination will also vary in accordance with the specific fluorescently labeled nucleotide assayed, e.g., variations in base, dye, or linker. Mutant DNA polymerase of the invention may exhibit anywhere from a slight reduction in discrimination against fluorescently labeled nucleotides to a complete elimination in discrimination, i.e., the mutant enzyme does not significantly differ with respect of rate of incorporation of labeled or unlabeled nucleotides. It is preferable to use embodiments of the subject mutant DNA polymerases that have at least a two-fold reduction in discrimination against one or more fluorescein type dye labeled nucleotides.

It will be appreciated by persons skilled in the art of molecular biology that the nucleotide label interaction region of a given DNA polymerase is defined with respect to a

specific fluorescently labeled nucleotide. Changes in one or more of the following parameters of the structure of a fluorescently labeled nucleotide may alter the identity of the amino acid residues that form the nucleotide label interaction site of a given DNA polymerase: (1) identity of the base, (2) the site of attachment on the nucleotide base, (3) the identity of the linker joining the base to the fluorescent dye, and (4) the identity of the fluorescent dye. The nucleotide labeled interaction region of Taq defined with respect to TET(II)•ddCTP comprises the amino acid residues E520, A531, L522, R523, E524, A525, H526, P527, I528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832. The sites at R660, T664, and E681 are of preferred sites for introducing mutations. Given that the 3-dimensional structure of Taq DNA polymerase (and other DNA polymerases) is well known and the three dimensional structure of TET(II)•ddCTP is understood with a high degree of certainty, the location of the amino acid residues that constitute the labeled nucleotide interaction region with respect to TET(II)•ddCTP may be translated to a different set of amino acid residues to accommodate structural differences between TET(II)•ddCTP and other fluorescently labeled nucleotides so as to define the labeled nucleotide interaction site with respect to those other nucleotides. For example, increasing the length of the linker between the base and the fluorescent label and the base may predictably alter the identity of amino acid residues that form the labeled nucleotide interaction site, even though the base, base attachment site, and fluorescent dye are the same. In many embodiments of the subject polymerases, the set of amino acid residues that form the labeled nucleotide interaction site with respect to a given fluorescently labeled nucleotide will overlap with the set of amino acid residues that form the labeled nucleotide interaction site as defined with respect to a second fluorescently labeled nucleotide.

Embodiments of the invention include mutant DNA polymerases that exhibit reduced discrimination against nucleotides labeled with fluorescein-type dyes, wherein the fluorescein type dye is joined to the nucleotide base by an alkynylamino-type linker. The fluorescein-type dye may be a fluorescent energy transfer dye, comprising a fluorescein-type dye moiety as a component of the energy transfer dye. In addition to reduced discrimination against fluorescently labeled nucleotides comprising an alkynylamino-type linker, the mutant DNA polymerases of the invention may also exhibit reduced discrimination against nucleotides

comprising other types of linker. In order to minimize steric interference between the polynucleotide and the fluorescent label, purines are usually labeled at position 7 and pyrimidines are usually labeled at position 5.

5 Mutant DNA polymerases of the invention have one or more discrimination reducing mutations at amino acid residue positions within the nucleotide label interaction region of a given DNA polymerase. Discrimination reducing mutations are usually, although not necessarily, substitution mutations. Several different amino residues may be substituted at a given position of a parent enzymes so as to give rise to a discrimination reducing mutations. The amino acid residues at a given residue position within the nucleotide label interaction
10 region may be systematically varied so as to determine which amino acid substitutions result in the reduction of discrimination against the fluorescein-type dye labeled nucleotide dye of interest and the degree of such a reduction in discrimination. The extent to which a particular mutation (or set of mutations) reduces discrimination may be measured by a selectivity assay as described in example 2. The substitution mutation is preferably, although not necessarily,
15 a mutation that reduces the size of the amino acid residue side chain of the amino acid residue present in the parent DNA polymerase. Mutations are preferably, although not necessarily, conservative so as to maintain the specific polar or non-polar character of the amino acid residue at the analogous position parent molecule. The mutations in the nucleotide label interaction region of a DNA polymerase preferably result in the substitution of the amino acid
20 residue of the parent enzyme with the amino acid residue at the corresponding position of phage T7 DNA polymerase (provided that a difference exists between the amino acid residues at that position in T7 polymerase and the parent enzyme).

Discrimination reducing mutations are in the nucleotide label interaction region of DNA polymerases. The nucleotide label interaction region is formed by portions of three
25 regions of the DNA polymerase. These three regions are located in (i) the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. Positions in Taq DNA polymerase that form the nucleotide label interaction region are positions E520, A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593,
30 R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832. Analogous positions in DNA polymerases other than Taq are also form

a nucleotide label interaction region. Preferred positions for substitution mutations are R595, D655, R660, and E681. A particularly preferred position for mutations is E681, with the preferred substitution at position 681 being M. Other suitable substitution mutations at E681 are as follows (listed in order of decreasing preference, except where note by an equal sign to denote approximate equivalence"): M>I>W>L>V>P>H=K=G=T=S>D=A=N>Y=C. A preferred substitution mutation at position R660 is R660D.

The specific amino acid residues that form the nucleotide interaction region will vary in accordance with the particular DNA polymerase selected as a parent enzyme for the introduction of discrimination reducing mutations. The determination of analogous amino acid residues positions between different DNA polymerases may easily be achieved by the person skilled in the art because of the large number of DNA polymerase amino acid sequences that have been determined and the many regions of homology have been found between these different DNA polymerases. For example, a large compilation of the amino acid sequences of DNA polymerases from a wide range of organism and homology alignments between the sequences can be found in Braithwaite and Ito, Nucl. Acids Res. 21(4):787-802 (1993). Examples of amino acid residues within the nucleotide label interaction regions of phage T7 polymerase and *E. coli* DNA polymerase are provided in Table 1. In addition to providing mutant DNA polymerases having reduced discrimination for fluorescein type dyes in Taq, T7 and *E. coli* DNA polymerase I, the invention provides mutant DNA polymerases from many other organisms. In general, the teachings of the invention may be used to produce mutant DNA polymerases having reduced discrimination for fluorescein type dyes from any DNA polymerase that shares sufficient amino acid sequence homology to Taq DNA polymerase to permit a person of ordinary skill in the art to identify one or more amino acid residue positions in the DNA polymerase that are analogous to positions E520, A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, E537, R573, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, L817, E820, L828, K831, and E832 in Taq DNA polymerase. Parent DNA polymerases that may be modified to contain discrimination reducing mutations in the nucleotide label interaction region include, but are not limited to, DNA polymerases from organisms such as *Thermus flavus*, *Pyrococcus furiosus*, *Thermotoga neapolitana*, *Thermococcus litoralis*, *Sulfolobus solfataricus*, *Thermatoga maritima*, *E. coli* phage T5, and *E. coli* phage T. The DNA

polymerases may be thermostable or not thermostable. It will be appreciated that the present invention enables persons skilled in the art to introduce fluorescein-type dye discrimination reducing mutations in to DNA polymerases from a wide variety of organisms, including DNA polymerases that have not been isolated at the time of the filing of this application provided.

5 Additionally, embodiments of the invention includes some purified naturally-occurring DNA polymerases that have the desired low degree of discrimination against fluorescently labeled nucleotides. Such naturally-occurring DNA polymerases are structurally and functionally analogous to the mutant DNA polymerases explicitly described herein.

10 The amino acid residues that constitute the nucleotide label interaction region of a given DNA polymerase vary in accordance with the specific fluorescently labeled nucleotide that is used to define the nucleotide label interaction region. Similarly, the mutations that are discrimination reducing mutations may vary in accordance with the specific fluorescently labeled nucleotide that is used to define the labeled nucleotide interaction region. Additionally, the degree of discrimination reduction achieved by the mutation (or mutations)
15 in the labeled nucleotide interaction site may vary with the specific labeled nucleotide of interest. For example, E681M is the preferred discrimination reducing mutation in Taq with respect to TET(II)•ddCTP resulting in a 47x reduction in discrimination and a significantly lower reduction in discrimination against a second fluorescently labeled nucleotide. Conversely, an E681T mutation may result in a high level reduction in discrimination against
20 the second fluorescently labeled nucleotide and only a low level of reduction in discrimination against TET(II)•ddCTP.

Given that a mutant DNA polymerase of the invention may have discrimination reducing mutation in the nucleotide label interaction region resulting in a significant degree of reduction in discrimination for a specific fluorescently labeled nucleotide and little or no
25 reduction in the degree of reduction of discrimination against another fluorescently labeled nucleotide (assuming there is significant discrimination against that fluorescently labeled nucleotide by the parent DNA polymerase), a given mutant DNA polymerase may be said to be "receptive" with respect to one or more given fluorescently labeled nucleotide. A specific mutant DNA polymerase is referred to as "receptive" with respect to a specific fluorescently
30 labeled nucleotide if a discrimination reducing mutation in the nucleotide label interaction site in the specific enzyme of interest results in at least a five fold reduction in discrimination against that given fluorescently labeled nucleotide. A mutant DNA polymerase of the

invention may be receptive with respect to more than one fluorescently labeled nucleotide. Conversely, a specific fluorescently labeled nucleotide may be "receptive" with respect to a given mutant DNA polymerase of the invention.

5 In embodiments of the subject mutant DNA polymerases comprising more than one discrimination reducing mutation in the nucleotide label interaction region, the mutation site may be in the same or different region of the three regions of a polymerase that form the nucleotide label interaction region. In general, mutant DNA polymerases of the invention will have 1, 2, or 3 discrimination reducing mutations. However, the invention also provides mutant DNA polymerases having more than 3 discrimination reducing mutations. By
10 combining multiple discrimination reducing mutations, greater levels of reduction in labeled nucleotide discrimination may be achieved. However, in many embodiments of the invention, mutant DNA polymerases have levels of reduced labeled nucleotide discrimination that are the same or less than the levels of DNA polymerase with single discrimination reduction mutations in the nucleotide label interaction region. Preferred combinations of mutations in
15 a Taq DNA polymerase background are R660D, E681G, and F667Y, i.e., Taq DNA polymerase mutant (R660D, E681G, and F667Y).

Different embodiments of DNA polymerase having mutations in the nucleotide label interaction region differ with respect to the degree of reduction in discrimination against specific fluorescently labeled nucleotides. These differences may be measured by an assay
20 in order to determine which specific embodiments have the greatest degree of reduction in discrimination against the particular fluorescently labeled nucleotides of interest. Generally, such assays measure competition between a fluorescently labeled nucleotide and an unlabeled nucleotide for incorporation into the same site on a primed template. One example of such an assay (referred to herein as a "selectivity assay") is described in detail below in Example
25 2.

The mutant DNA polymerases of the invention may comprise numerous mutations in addition to discrimination reduction mutations in the nucleotide label interaction region. These secondary mutations may be either inside or outside the nucleotide label interaction region. Secondary mutations may be selected so as to have as to confer some useful property
30 on the mutant DNA polymerase. For example, additional mutations may be introduced to increase thermostability; decrease thermostability, increase processivity, decrease processivity, decrease 3'-5' exonuclease activity, increase 3'-5' exonuclease activity, decrease 5'-3'

exonuclease activity, increase 5'-3' exonuclease activity, and increase incorporation of dideoxynucleotides. Alternatively, the secondary mutations may be essentially neutral in known effect.

5 Of particular interest are embodiments of the subject mutant DNA polymerase that comprise one or more secondary mutation that reduce 3'-5' exonuclease activity. DNA polymerases that are deficient in 3'-5' exonuclease activity have superior properties for PCR and for chain termination polynucleotide sequencing. Mutations that reduce 3'-5' exonuclease activity in DNA polymerase are well known to person of ordinary skill in the art. Detailed guidance on how to introduce mutations that reduce 3'-5' exonuclease activity can
10 be found, among other places in U.S. Patent No. 4,795,699 (Tabor); U.S. Patent No. 5,541,099; U.S. Patent No. 5,489,523; and Bernad et al., *Cell* 59:219-288 (1989). Examples of such mutations in Taq DNA polymerase include G46D. For embodiments of the mutant DNA polymerases that are used for sequencing, it is preferable to include a G46D (or analogous mutations in DNA polymerases other than Taq) in addition to mutations in the
15 nucleotide label interaction region.

Also of interest among secondary mutations in the subject DNA polymerase mutants are mutations that increase incorporation of dideoxynucleotides, i.e., reduce the ability of
a DNA polymerase to discriminate against dideoxynucleotide as opposed to deoxynucleotides. Guidance on making such mutations can be found, among other places in
20 published PCT application WO96/12042 (application number PCT/US95/12928). Of particular interest is the mutation F667Y in Taq and analogous mutations in other DNA polymerase. While F667Y is not part of the nucleotide label interaction region in Taq DNA polymerase with respect to Tet(II)-ddLTP, F667Y mutations may reduce discrimination against fluorescein-type dye labeled nucleotides (see Table 1). Accordingly, for use in certain
25 procedures, e.g., DNA sequencing, be desirable to combine an F667Y mutations with one or more discrimination reducing mutations in the nucleotide label interaction region so as to reduce discrimination of the polymerase between deoxynucleotides and 2'3' dideoxynucleotides. Mutant DNA polymerase of the invention having the F667Y mutation (or equivalent thereof) are particularly useful in Sanger type DNA sequencing with
30 fluorescently labeled 2'3' dideoxynucleotide chain terminators.

Numerous genes encoding DNA polymerases have been isolated and sequenced. This sequence information is available on publicly accessible DNA sequence databases such as

GENBANK. A large compilation of the amino acid sequences of DNA polymerases from a wide range of organism can be found in Braithwaite and Ito, Nucl. Acids Res. 21(4):787-802 (1993). This information may be used in designing various embodiments of DNA polymerases of the invention and polynucleotide encoding these enzymes. The publicly available sequence information may also be used to clone genes encoding DNA polymerases through techniques such as genetic library screening with hybridization probes.

Other embodiments of the invention are polynucleotide sequences encoding the mutant DNA polymerases provided herein. Polynucleotide sequences encoding the mutant DNA polymerase of the invention may be used for the recombinant production of the mutant DNA polymerases. Polynucleotide sequences encoding mutant DNA polymerases having reduced discrimination against fluorescently labeled nucleotide may be produced by a variety of methods. A preferred method of producing polynucleotide sequences encoding mutant DNA polymerases having reduced discrimination against fluorescently labeled nucleotides is by using site-directed mutagenesis to introduce desired discrimination reducing mutations into polynucleotides encoding the parent DNA polymerase molecules. Site-directed mutagenesis techniques are well known in the art as exemplified by U.S. Patent No. 4,711,848; U.S. Patent No. 4,873,192; U.S. Patent No. 5,071,743; U.S. patent, 5,284,760; U.S. Patent No. 5,354,670; U.S. Patent No. 5,556,747; Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982), and Edelman et al DNA 2:183 (1983). Detailed protocols for site-directed mutagenesis are also given many general molecular biology textbooks such as Sambrook et al Molecular Cloning a Laboratory Manual 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor (1989), Ausubel et al. Current Protocols in Molecular Biology, (current edition). Additionally, many text books on PCR (the polymerase chain reaction), such as Diefenbach and Dveksler, PCR Primer: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY (1995), describe methods of using PCR to introduce directed mutations. Genes encoding parent DNA polymerase may be isolated using conventional cloning techniques in conjunction with publicly-available sequence information. Alternatively, many cloned polynucleotide sequences encoding DNA polymerases have been deposited with publicly-accessible collection sites, e.g., the American type culture collection deposit accession number ATCC 40336 is a phage clone of Taq DNA polymerase.

In addition to producing the mutant DNA polymerase encoding polynucleotides of the invention by introducing directed mutations into polynucleotides encoding parent DNA

polymerases, it is possible (although difficult) to produce the polynucleotides of the invention primarily by *in vitro* DNA synthesis techniques. *In vitro* DNA synthesis techniques are well known to those skilled in the art and examples of *in vitro* DNA synthesis can be found in U.S. Patent No. 5,252,530; U.S. Patent No. 4,973,679; U.S. Patent No. 5,153,319; U.S. Patent No. 4,668,777; U.S. Patent No. 4,500,707; U.S. Patent No. 5,132,418; U.S. Patent No. 4,415,732; U.S. Patent No. 4,458,066; and U.S. Patent No. 4,811,218. When producing relative polynucleotide molecules by *in vitro* DNA synthesis, smaller molecules are usually produced first and subsequently joined together by hybridization and ligation. Mutant DNA polymerase encoding polynucleotides may also be produced by a combination of *in vitro* synthesis and site-directed mutagenesis of cloned genes.

Polynucleotide encoding the mutant DNA polymerase of the invention may be used for the recombinant expression of the mutant DNA polymerases. Generally, the recombinant expression of the mutant DNA polymerase is effected by introducing a mutant DNA polymerase into an expression vector adapted for use in particular type of host cell. Thus, another aspect of the invention is to provide expression vectors comprising a polynucleotide encoding a mutant DNA polymerase of the invention, such that the polymerase encoding polynucleotide is functionally inserted into the expression vector. The invention also provides host cells comprising the expression vectors of the invention. Host cells for recombinant expression may be prokaryotic or eukaryotic. Examples of host cells include bacterial cells, yeast cells, cultured insect cell lines, and cultured mammalian cell lines. Preferably, the recombinant host cell system is selected so as to closely match the organism from which the mutant DNA polymerase was derived. For example, prokaryotic DNA polymerases are preferably expressed in a prokaryotic expression system. A wide range of expression vectors are well known in the art. Description of various expression vectors and how to use them can be found among other places in U.S. Patent No. 5,604,118; U.S. Patent No. 5,583,023; U.S. Patent No. 5,432,082; U.S. Patent No. 5,266,490; U.S. Patent No. 5,063,158; U.S. Patent No. 4,966,841; U.S. Patent No. 4,806,472; U.S. Patent No. 4,801,537; and Goeddel et al., Gene Expression Technology, Methods of Enzymology, Vol. 185, Academic Press, San Diego (1989). The expression of DNA polymerases in recombinant cell systems is a well-established technique. Examples of the recombinant expression of DNA polymerase can be found in U.S. Patent No. 5,602,756; U.S. Patent No. 5,545,552; U.S. Patent No. 5,541,311; U.S. Statutory Inventor Registration H1,531; U.S. Patent No. 5,500,363; U.S. Patent No. 5,489,523; U.S.

Patent No. 5,455,170; U.S. Patent No. 5,352,778; U.S. Patent No. 5,322,785; and U.S. Patent No. 4,935,361.

Other embodiments of the invention include multiple DNA polymerase compositions particularly useful for polynucleotide sequencing, such compositions comprise at least two different mutant DNA polymerases of the invention, wherein (1) the first mutant DNA polymerase is receptive with respect to a first fluorescently labeled nucleotide; (2) the second mutant DNA polymerase is receptive with respect to a second fluorescently labeled nucleotide; and (3) the first and second fluorescently labeled nucleotides differ from one another with respect to their nucleotide bases and fluorescent labels. The first and second fluorescently labeled bases may also differ with respect to one another by way of the linker, the base attachment position, or the fluorescent dye attachment site. The subject compositions are useful for catalyzing the sequencing reactions in Sanger type DNA sequencing with fluorescent dye labeled 2'3' dideoxy chain terminating nucleotides. Chain termination sequencing with fluorescently labeled terminators preferably employs at least two, and more preferably 4 different fluorescently labeled chain terminators, wherein each different base is labeled with a distinctive fluorescent label. Because of the necessary structural differences between the different fluorescently labeled chain terminators required for a sequencing reactions, i.e., nucleotide bases and fluorescent labels, there are many mutant DNA polymerases of the invention that are not receptive to all of the fluorescently labeled terminators necessary for a given sequencing reaction. Thus, there are embodiments of the subject DNA polymerases that may have undesirably high levels of discrimination against one or more of the labeled terminators used in a sequencing reaction set. The subject compositions of two or more mutant polymerases ameliorates this problem by simultaneously employing multiple mutant DNA polymerases that are receptive to different chain labeled terminators, thereby having at least one of the mutant polymerases "compensate" for the discrimination against a particular fluorescently labeled terminator by the other polymerases catalyzing the sequencing reactions. The ratio of the different DNA polymerases in the composition preferably are selected so as to result in approximately equal levels of total activity for each of the different mutant DNA polymerases. Differences in specific activity between the different mutant polymerases may be taken into account when equalizing total activity ratios between the polymerases. Differences in activity levels between the various mutant DNA polymerases in the subject compositions may also be compensated for by

adjusting the levels of the different fluorescently labeled terminators in the subject compositions. The subject multiple polymerase compositions may comprise two, three, four, or more different mutant DNA polymerases. The mutant polymerase may or may not be derived from the same species or strain. The different mutation DNA polymerases in the subject mutant polymerase compositions may or may not be receptive for one or more of the fluorescently labeled nucleotides in a given set fluorescently labeled dideoxynucleotides for sequencing.

The invention also includes various methods of using the mutant DNA polymerases (or subject multiple mutant DNA polymerase compositions) of the invention. The mutant DNA polymerases of the invention may be substituted for the corresponding parent DNA polymerases in most procedures that employ DNA polymerases. In order to more fully take advantage of the properties of the subject mutant DNA polymerases, the amount (or concentration) of labeled and unlabeled nucleotides used in the methods of the invention may be changed with respect to the amounts (or concentrations) used in the corresponding methods employing convention DNA polymerases. These changes in the amount of nucleotide may be optimized by routine experimentation. Methods of the invention comprise the step of extending a primed polynucleotide template with at least one fluorescently labeled nucleotide, wherein the extension is catalyzed by a mutant DNA polymerase of the invention. Thus, the subject methods result in the formation of one or more different fluorescently labeled polynucleotides produced by primer extension. The subject methods of synthesizing a fluorescently labeled polynucleotide may be used in a variety of procedures including, but not limited to, Sanger sequencing (e.g., dideoxy nucleotide chain termination), the polymerase chain reaction (PCR), polynucleotide labeling, minisequencing. The reduced discrimination against fluorescently labeled nucleotide properties of the subject mutant DNA polymerase is particularly useful for Sanger DNA sequencing reactions, including cycle sequencing. The use of the subject mutant DNA polymerases for Sanger sequencing reduces the amount of fluorescently labeled chain terminating nucleotides required for a sequencing reaction and may in many cases be used to increase the number of bases that may be identified in single sequencing reaction that is analyzed on an automated fluorescence-based sequencing apparatus such as an Applied Biosystems 310 or 377 (Applied Biosystems Division of Perkin-Elmer, Foster City, CA.). Detailed protocols for Sanger sequencing are known to those skilled in the art and may be found, for example in Sambrook et al, Molecular Cloning, A

Laboratory Manual, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).

The invention also provides kits for synthesizing fluorescently labeled polynucleotides. The kits may be adapted for performing specific polynucleotide synthesis procedures such as DNA sequencing or PCR. Kits of the invention comprise a mutant DNA polymerase of the invention and a fluorescently labeled nucleotide that exhibits reduced discrimination with respect to the mutant DNA polymerase in the kit. Kits preferably contain detailed instructions on how to perform the procedures for which the kits are adapted. Optionally, the subject kit may further comprise at least one other reagent required for performing the method the kit is adapted to perform. Examples of such additional reagents include unlabeled nucleotides, buffers, cloning vectors, restriction endonucleases, sequencing primers, and amplification primers. The reagents include in the kits of the invention may be supplied in premeasured units so as to provide for greater precision and accuracy.

Other embodiments of the invention include kits comprising (1) the subject compositions of multiple mutant DNA polymerases, and (2) fluorescently labeled chain terminating nucleotides suitable for use with the subject compositions, i.e., each labeled chain terminator is receptive with respect to at least one of the mutant DNA polymerases in the composition. Additional embodiments of the invention include kits for sequencing DNA that comprise a multiple mutant polymerase composition of the invention and at least two different fluorescently labeled chain terminating nucleotides are labeled at different bases, wherein each of the fluorescently labeled chain terminating nucleotides is receptive with respect to at least one mutant DNA polymerase in the composition.

The invention, having been described above, may be better understood by reference to the following examples. The examples are offered, for among other reasons, to illustrate specific embodiment of the invention and should not be construed as a limitation on the invention.

EXAMPLES

Example 1

Purification of Mutant Forms of Taq DNA Polymerase

Lysates of *E. coli* containing recombinant constructs designed for the production of

recombinant mutant Taq DNA polymerases were made essentially as described in tDesai, U.J. and Pfaffle, P.K., *Biotechniques*, 19:780-784 (1995). In order to prevent the polymerase from binding to chromosomal and plasmid DNAs contaminating the lysate, 5 M NaCl was added dropwise to the heat treated, clarified lysates to bring the final NaCl concentration to 0.25 M. DNA was then precipitated from this mixture by dropwise addition of 5% polyethylimine (in 20 mM TRIS•Cl, pH 8.5) to make the final concentration of PEI 0.3%. Precipitation was allowed to continue for 5 minutes on ice. A white, cloudy precipitate was removed by centrifugation at 15,000 x g for 15 minutes at 4°C. The supernatant fluid was decanted and saved. Following centrifugation, the NaCl concentration was reduced to 0.13 M by monitoring conductivity of the solution during the addition of TETT minus NaCl (20 mM TRIS•Cl, 0.1 mM EDTA, 0.05% Tween-20, 0.05% Triton-X100, 1% glycerol, pH 8.5).

Excess PEI was removed using a Bio-Rex 70 (BIO-RAD, Richmond, CA) column (2.5 x 30 cm). The column was poured and equilibrated with TETT Buffer + 0.1 M NaCl. The polymerase does not bind to the Bio-Rex 70 under these conditions.

To remove contaminating *E. coli* proteins, the Bio-Rex 70 column eluate was loaded directly onto a Heparin-Agarose (Sigma Chemical Company, St. Louis, MO) column (1.5 x 30 cm) which was also poured and equilibrated in TETT Buffer + 0.1 M NaCl. The heparin-agarose column was washed with 2 column volumes of TETT + 0.1 M NaCl and Taq DNA polymerase was eluted as a sharp peak using TETT + 1 M NaCl. Elution was monitored at 280 nm.

The heparin-agarose column fractions corresponding to the peak absorbance were pooled and concentrated to 0.15 ml using Ultrafree-15 Centrifugal Filter Devices (Millipore Corporation, MA) according to the manufacture's recommendations for centrifugation speeds and times. The concentrate was diluted to 15 ml with TETT Buffer + 5% glycerol and the sample was re-concentrated to 0.15 ml. This was repeated one more time to reduce the final NaCl concentration below 1 mM in the protein samples.

The concentrated polymerase samples were diluted two-fold using TETT + 5% glycerol and an equal volume of TETT + 95% glycerol was added to bring the final glycerol concentration to about 50%. Samples were stored at -20 °C. Protein concentrations were determined using the "Bradford Protein Assay" (BIO-RAD, Richmond, CA). Activity was measured using a radiometric assay (described elsewhere).

Typical yields of polymerase from 2-liters of induced *E. coli* culture (corresponding

to 30- 50 ml of heat treated, clarified lysate) ranged from 4 to 24 mg. SDS-PAGE analysis of the purified samples showed one dark band of about 94,000 molecular weight) and several minor ones after Coomassie Blue staining. The gels indicated a typical purification level of > 90%.

5

Example 2

Selectivity Assay

An unlabeled versus dye-labeled terminator assay ("terminator" is defined as a non-extendible base such as 2',3'- ddNTPs) was used to screen mutant Taq DNA polymerase samples for better Tet(II)-ddCTP incorporating mutant forms of this polymerase. This assay is based upon two substrates competing for the same active site at the same time during a steady state reaction in which only the polymerase concentration is limiting. Therefore, the assay measure the polymerase's "selectivity" for the unlabeled versus the fluorescein-labeled terminator. The DNA Primer/template used in this assay format is given below:

15

5'->(FAM)-CCC TCG CAG CCG TCC AAC CAA CTC A

GGG AGC GTC GGC AGG TTG GTT GAG **TGC** CTC TTG TTT<-5'

The next template position following the 3'-end of the primer is indicated above by the bold and underlined G.

20

The reaction consisted of:

80 mM TRIS•Cl (pH 9.0 at 20 °C)

1000 nM DNA primer/template [5'-(FAM)25mer / 36 G₁ template]

2 mM MgCl₂

25

50 μM TET(II)-ddCTP

1 μM ddCTP

0.25 Units of enzyme

40 μL reaction volume

60 °C reaction temperature

30

Samples (2 μL) were removed from the reaction mixture at pre-determined times (typically, 20 second intervals for 0.25 Units of polymerase activity per μL) and added to ice

cold 50 μ L 0.5 M EDTA (pH 8.0). Timed aliquots were mixed and held on ice for further processing.

Samples of each time point were processed to remove excess, unincorporated TET(II)•ddCTP. Typically, 1.6 μ L of each quenched sample were added to 250 μ L of 0.8 M LiCl plus 0.2 μ g/ml *E. coli* tRNA, followed by 750 μ L of 95% ethanol. After mixing, the nucleic acids were allowed to precipitate for 20 minutes at -20 °C. The precipitates were recovered by centrifugation using standard procedures. The supernatant fluid was discarded and pellets were dissolved in 50 μ L of 50% formamide. Gel samples were heat treated (95°C for 2 minutes) and 2 μ L were loaded per sample lane on a 16% denaturing DNA sequencing gel. Gels were run on an Applied Biosystems Model 373 Sequencer using GeneScan Fragment Analysis software to measure the amount of FAM fluorescence in the bands corresponding to the 25-mer primer, the 26-mer product (indicating a ddC incorporation event) and the apparent "27-mer" product band (indicating a TET(II)•ddC incorporation event).

The fluorescence signal in each of the bands was summed and the percent of signal in each band was used for further calculations as a normalization to avoid lane to lane loading differences. Energy transfer from the %-FAM moiety present on the apparent "27-mer" product molecules to the Tet(II) moiety on the newly incorporated 3'-base was not corrected since all ratios were compared to "wild type" or Taq G46D.) The normalized fluorescent signals in the 26-mer and "apparent" 27-mer product bands were corrected for the different concentrations of the two molecules used in the reaction and the corrected values were plotted versus time. The velocity of incorporation for each substrate was determined using least square fits to the data. The ratio of ddC / TET(II)•ddC incorporation rates is equal to the selectivity bias that the sample polymerase shows for the unlabeled versus the TET(II)-labeled nucleotides and reflects the following relationship:

$$\frac{V_{ddC}}{V_{Tet(II) \cdot ddC}} = \frac{(k_{cat} / K_M)_{ddC} [ddC]}{(k_{cat} / K_M)_{Tet(II) \cdot ddC} [Tet(II) \cdot ddC]}$$

where:

$$\begin{aligned} V_{ddC} &= \text{velocity of ddC incorporation} \\ V_{Tet(II) \cdot ddC} &= \text{velocity of Tet(II)•ddC incorporation} \\ k_{cat} &= \text{catalytic rate constant} \end{aligned}$$

K_M = nucleotide equilibrium binding constant
 $[ddC]$ = concentration of ddCTP in the reaction
 $[Tet(II) \cdot ddC]$ = concentration of Tet(II)•ddCTP in the reaction

5 In this assay format, "wild-type" Taq or (Taq G46D) showed a selectivity bias or ddC / Tet(II)•ddC number of about 85 to 1. Mutants showing lower selectivity bias ratios were submitted to further testing. The Table 2 below shows the results for a few of the mutants tested by way of a few examples:

10

Table 2

<u>Taq</u>	<u>Selectivity Number</u>	<u>WT / Mutant</u>
G46D	85	85 / 85 or 1
G46D; R660D	8	85 / 8 or ≈ 10
15 G46D; R595E	28	85 / 28 or ≈ 3
G46D; F667Y	28	85 / 28 or ≈ 3
G46D; E681G	40	85 / 40 or ≈ 2
G46D; D655L	40	85 / 40 or ≈ 2

20

Example 3Next Nucleotide Rate Effect Assay

25 An additional kinetic step between "ground state" nucleotide binding or initial collision and correct base pair formation and the group transfer reaction would be expected to slow the polymerase dissociation rate from an Enz•DNA complex having a 3'-dideoxynucleotide in an assay termed the "Next Nucleotide Rate Effect" (Patel et al., 1991). This assay measures the steady state rate of incorporation of ddTTP (i.e., the enzyme is limiting) in the absence or presence of the next correct nucleotide. The primer template pair is shown below:

30 5'-(FAM)-CCC TCG CAG CCG TCC AAC CAA CTC A
 GGG AGC GTC GGC AGG TTG GTT GAG **T**AG GTC TTG TTT<-5'

The next template position is indicated by the bold, underlined A. The next template

position beyond A is G. Under steady state reaction conditions, essentially all of the available polymerase is bound to the primer/template. When ddTTP is present alone in solution, it is incorporated following binding to its template position, A. Additional incorporation events require the polymerase to dissociate from the Enz•DNA complex and find another available primer/template that has not already undergone and incorporation event. Hence the rate of incorporation under these conditions is the dissociation rate of the polymerase from the Enz•DNA complex. If the next correct nucleotide, dGTP or ddCTP, is also present in the reaction mixture, the dissociation rate of the polymerase from the Enz•DNA•ddCTP complex, for example, will be slower if there is an additional kinetic step between the group transfer reaction that incorporated the ddTTP and an attempt by the polymerase to incorporate ddCTP in a processive mode of synthesis. This slower rate of dissociation can be detected as a slower incorporation rate of ddTTP since no chemistry can occur once ddTTP and the polymerase can no longer be processive despite the presence of another correct nucleotide. As shown in Figure 2, the presence of the next correct nucleotide does indeed slow the turnover or dissociation rate of the polymerase (Taq G46D; F667Y). Figure 2 also shows that the presence of a fluorescein dye on the next correct nucleotide (in this case, Tet(II)•ddCTP), appears to accelerate the turnover rate. We interpret this to mean that the polymerase is constantly undergoing a conformational change and that it can attempt to undergo the change even in the absence of the next correct nucleotide. However, the presence of a fluorescein dye on the next correct nucleotide blocks the ability of the polymerase to undergo such a change and thereby causes an immediate dissociation of the enzyme following the group transfer step for ddTTP incorporation. Hence, the fluorescein dye appears to accelerate the polymerase dissociation rate by eliminating a kinetic step (or steps) following the group transfer reaction.

Figure 3 shows the results for a Next Nucleotide Rate Effect assay for a "multiple" mutant form of Taq DNA polymerase, Taq G46D; R660D; F667Y; E681G. In this case, the presence of Tet(II) on the next correct nucleotide is "transparent" to the mutant polymerase. We interpret this to mean that the mutant polymerase can indeed undergo the same kinetic steps following group transfer that "wild-type" versions of this polymerase undergo. We also interpret these results to indicate that the F667Y mutation belongs in a different class than the R660D or E681G mutations since Taq G46D; F667Y still shows a "fluorescein-effect" in the "Next Nucleotide Rate Effect" assay, however, the multiple mutant, Taq G46D; R660D, F667Y; E681G, does not.

Typical assay conditions for the Next Nucleotide Effect assay were as follows:

1000 nM primer/template DNA
80 mM TRIS•Cl (pH 9.0 @ 20° C)
2.4 mM MgCl₂
0.02 Units/μL polymerase activity
400 μM each nucleotide (when present)

Samples were taken and processed in the same manner as described under "Selectivity Assay." In this case, it is possible to distinguish a ddC-incorporation event from a Tet(II)•ddC incorporation event by the migration rate of the resulting fragments in a 16% gel. Incorporation of ddC results in a "normal" 26-mer band that migrates as expected above or slower than the 25-mer primer. Incorporation of Tet(II)•ddC results in slower migration causing the band to migrate with an apparent size equivalent to a 27- or 28-mer.

Example 4

Analysis of Additional Mutants

Table 1, provided below, provides a summary of results obtained with selectivity assays performed with several different Taq mutants. The analogous site for the mutation in the enzymes *E. coli* DNA polymerase I and phage T7 DNA polymerase are also noted. The term "FS" refers to a Taq DNA polymerase having a F667Y mutation.

References

- Barnes, W.M. (1992) The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion. *Gene* 112: 29-35.
- Brandis, J.W., Edwards, S.G. and Johnson, K.A. (1996) Slow rate of phosphodiester bond formations accounts for the strong bias that Taq DNA polymerase shows against 2',3'-dideoxynucleotide terminators. *Biochemistry* 35: 2189-2200.
- Desai, U.J. and Pfaffle, P.K. Single-step purification of a thermostable DNA polymerase expressed in *Escherichia coli*. *Biotechniques* 19: 780-784.

Fersht, A. (1985) in "Enzyme Structure and Function," W.H. Freeman and Company,
2nd ed., pp. 111-112.

Johnson, K.A. (1993) Conformational coupling in DNA polymerase fidelity.
5 *Ann. Rev. Biochem.* 62: 685-713.

Patel, S.S., Wong, I., and Johnson, K.A. (1991) Pre-steady-state kinetic analysis of
processive DNA replication including complete characterization of an exonuclease-
deficient mutant. *Biochemistry* 30: 511-525.

10

Incorporation by Reference

This application incorporates all publications, patents, and patent application
referenced herein in there entirety.

15

Equivalents

While the invention has been described and illustrated with reference to specific
embodiments, those skilled in the art will recognize that modifications and variations may be
made without departing from the principles of the invention as described hereinabove and set
forth in the following claims.

No.	Taq Mutant	Pol I Equiv.	T7 Equiv.	[Final] µg/µl	*Units per µl	Spec. Act.	TET(II)ddCTP/ ddCTP	ROX-ddCTP/ ddCTP	TAMRA-ddTTP/ ddTTP
							[Mut / Wt]	[Mut / Wt]	[Mut / Wt]
1	G46D			4.4	200	45.5	1*	1	1
	G46D (LS.1)			28.0	1950	69.6			
	G46D (LS.2)			78.3	6240	79.7			
	FS			RMS			3		920
2	G46D; F667W (LS.1)	F762	Y526	3.6	5	1.4	1	1	
		F762	Y526	25.0	150	6.0			
3	G46D; R573E	R668	R429	11.0	0	0.0			
4	G46D; E615L	E710	E480	4.8	0	0.0			
5	G46D; E615D	E710	E480	10.4	800	57.7	no activity		
6	G46D; E615I	E710	E480	7.0	140	20.0	no activity		
7	G46D; R587K	R682	*V443*	8.9	420	47.2	1	1	
8	G46D; R573K	R668	R429	9.3	0				
9	G46D; L657T	E752	T517	9.5	450	47.4	1	1	
10	G46D; R587K	R682	*V443*	nd	nd				
11	G46D; Q754S	Q849	Q615	13.0	0	0.0			
12	G46D; E615K	E710	E480	1.8	0	0.0			
13	G46D; R573Q	R668	R429	14.0	800	57.1	1	1	
14	G46D; D655L	T750	L515	11.3	400	35.4	2	0.7	
15	G46D; Q754K	Q849	Q615	nd	nd	nd			
16	G46D; R595K	R690	H460	nd	nd	nd			
17	G46D; K831M	H928	H704	8.3	300	38.1	1	1	
18	G46D; L682G	L777	*I540*	5.5	200	38.4	1	1	6
19	G46D; R659K	R754	D519	22.9	150	6.6	no activity		
20	G46D; A683E	N778	V641	12.8	900	70.3	1		
21	G46D; Q754K	Q849	Q615	5.0	0	0.0			
22	G46D; R593H	R688	E458	16.7	700	41.9	1		
23	G46D; R595E	R690	H460	23.5	50	2.1	3		
24	G46D; A683V	N778	V641	11.4	340	29.8	1		
25	G46D; Q592A	R687	A457	nd	nd	nd			
26	G46D; R660D	R755	D519	13.3	190	14.3	10		
27	G46D; T640G	R735	T507	13.0	225	17.3	1		
28	G46D; E681G	Q776	*I540*	7.5	170	22.7	2		
29	G46D; V654E	V749	E514	9.2	210	22.8			
30	G46D; Q613E	Q708	G478	15.2	71	4.7			
31	G46D; D610A	Q705	D475	16.6	0	0.0			
32	G46D; E820K	E917	E893	11.6	475	40.9			
33	G46D; L817A	L914	L890	15.2	470	30.9			
34	G46D; I684G	I779	G542						
35	G46D; R660D; F667Y			19.1	179	9.4	10		
36	G46D; R595D; R660D; F667Y			10.5	0	0.0			
37	G46D; D655L; R660D; F667Y			18.8	228	12.1	10		
38	G46D; R660D; F667Y; E681G			13.1	404	30.8	12		
39	G46D; R595E; F667Y			9.5	0	0.0			
40	G46D; T7 Loop JK			12.7	0	0.0			
41	G46D; A582Q583			13.2	0	0.0			
42	G46D; P656S	S751		16.2	560	34.6			
		K731							
40 lysates				22 act. mutants		17 tested 4 > "wt"			
						* WT = 85			

Table 1

R660- Mutants

	<u>Lysate #</u>	<u>Genotype</u>	<u>Specific Activity</u>	<u>Tet Selectivity</u> TET(II)-ddCTP / ddCTP * (Mutant / WT)
<u>Acidic-</u>				
Aspartic acid	29	CS; R660D	14	10*
	38	FS; R660D	9	10
	39	R595E; FS; R660D	0	nd
	40	D655L; FS; R660D	12	10
	41	FS; R660D; E681G	31	12
	49	CS; R660D	41	nd
Glutamic Acid	51	FS; R660E; E681G	11	7
	72	FS; R660E	1	.7
<u>Basic-</u>				
Lysine	50	FS; R660K	28	1**
Histidine	101	FS; R660H	13	1
<u>Imino-</u>				
Proline	66	FS; R660P	8	1
<u>Aliphatic-</u>				
Alanine	68	FS; R660A	4	4
Isoleucine	73	FS; R660I	5	0.9***
Valine	90	FS; R660V	10	1
	55	FS; R660V; E681G	1	1
Leucine	91	FS; G660L	8	0.6***
	52	FS; R660L; E681G	28	1
Glycine	47	FS; R660G; E681G	18	6
	78	FS; R660G	8	2
<u>Polar Uncharged-</u>				
Glutamine	53	CS; R660Q	47	1
	69	FS; R660Q	5	3
Serine	98	FS; R660S	16	7
Cysteine	93	FS; R660C	14	4
Asparagine	97	FS; R660N	13	3
Threonine	96	FS; R660T	26	3
Methionine				
<u>Aromatic-</u>				
Phenylalanine	92	FS; R660F	9	0.1***
Tyrosine	95	FS; R660Y	17	1

Table 2

E681- Mutants

	<u>Lysate #</u>	<u>Genotype</u>	<u>Specific Activity</u>	<u>Tet Selectivity</u> TET(II)•ddCTP / ddCTP * (Mutant / WT)
<u>Acidic-</u>				
Aspartic acid	71	FS; E681D	9	4**
<u>Basic-</u>				
Lysine	75	FS; E681K	52	6
Arginine				
Histidine	86	FS; E681H	37	7
<u>Imino-</u>				
Proline	74	FS; E681P	19	9
<u>Aliphatic-</u>				
Alanine	63	FS; E681A	13	6
Isoleucine	99	FS; E681I	37	27
Valine	76	FS; E681V	110	10
Leucine	87	FS; E681L	22	14
Glycine	48	FS; E681G	37	6
<u>Polar Uncharged-</u>				
Glutamine				
Serine	61	FS; E681S	12	5
Cysteine	88	FS; E681C	20	2
Asparagine	89	FS; E681N	40	4
Threonine	81	FS; E681T	35	6
Methionine	85	FS; E681M	32	47
<u>Aromatic-</u>				
Phenylalanine				
Tyrosine	80	FS; E681Y	42	3
Tryptophan	84	FS; E681W	37	17

11-05-97

*Ratio > 1 means improved TET(II)•ddCTP incorporation.
enzyme.

**Ratio = 1 means wild-type activity.

***Ratio < 1 means activity worse than wild-type.

M > I > W > L > V > P > H = K = G = T = S > D = A = N > Y = C
47 27 17 14 10 9 7 6 6 6 5 4 4 4 3 2

Table 2 (Continued)

Tryptophan 94

FS; R660W

8

1

D > E=S > C=A=Q=T=N> G > K=P=V=Y=W=H>I = L >> F
 10 7 7 4 4 3 3 3 2 1 1 1 1 1 0.9 0.6 0.1

*Ratio > 1 means improved TET(II)-ddCTP incorporation. Must be "85" to be "transparent to the enzyme."

**Ratio = 1 means wild-type activity.

***Ratio < 1 means activity worse than wild-type.

Table 2 (Continued)

CLAIMS

What is claimed is:

- 5 1. A DNA polymerase having at least one mutation in the nucleotide label interaction region, wherein the DNA polymerase has reduced discrimination for fluorescein-type dye labeled nucleotides.
- 10 2. A DNA polymerase according to claim 1 , wherein the mutation is in portion of the nucleotide-label interaction region selected from the group consisting of (i)the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop.
- 15 3. A DNA polymerase according to claim 2, wherein the mutation is segment of the enzyme corresponding to amino acid residue selected from the group consisting of E520, A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832.
- 20 4. A DNA polymerase according to claim 3, wherein the mutation is at a position selected from the group consisting of, R595, D655, R660, T664 and E681.
- 25 5. A DNA polymerase according to claim 4, wherein the DNA polymerase is Taq DNA polymerase.
6. A DNA polymerase according to claim 5, wherein the mutation is selected from the group consisting of R660D, D655L, E618G, and R595E.
- 30 7. A DNA polymerase according to claim 6, wherein comprising a mutation set belonging to the group consisting of (G46D, R660D, F667Y), (G46D, R595D, R660D, F667Y), and (G46D, R660D, F667Y, E681G), and (G46D, F667Y, E681G).

8. A DNA polymerase according to claim 2, wherein the DNA polymerase is a thermostable DNA polymerase.

9. A polynucleotide encoding a DNA polymerase according to claim 1.

5

10. An expression vector having a promoter, wherein the vector comprises a polynucleotide according to claim 1 in functional combination with the promoter.

11. A host cell comprising an expression vector according to claim 10.

10

12. A method of synthesizing a fluorescently labeled polynucleotide, said method comprising the step of mixing a DNA polymerase according to claim 1 with a primed template.

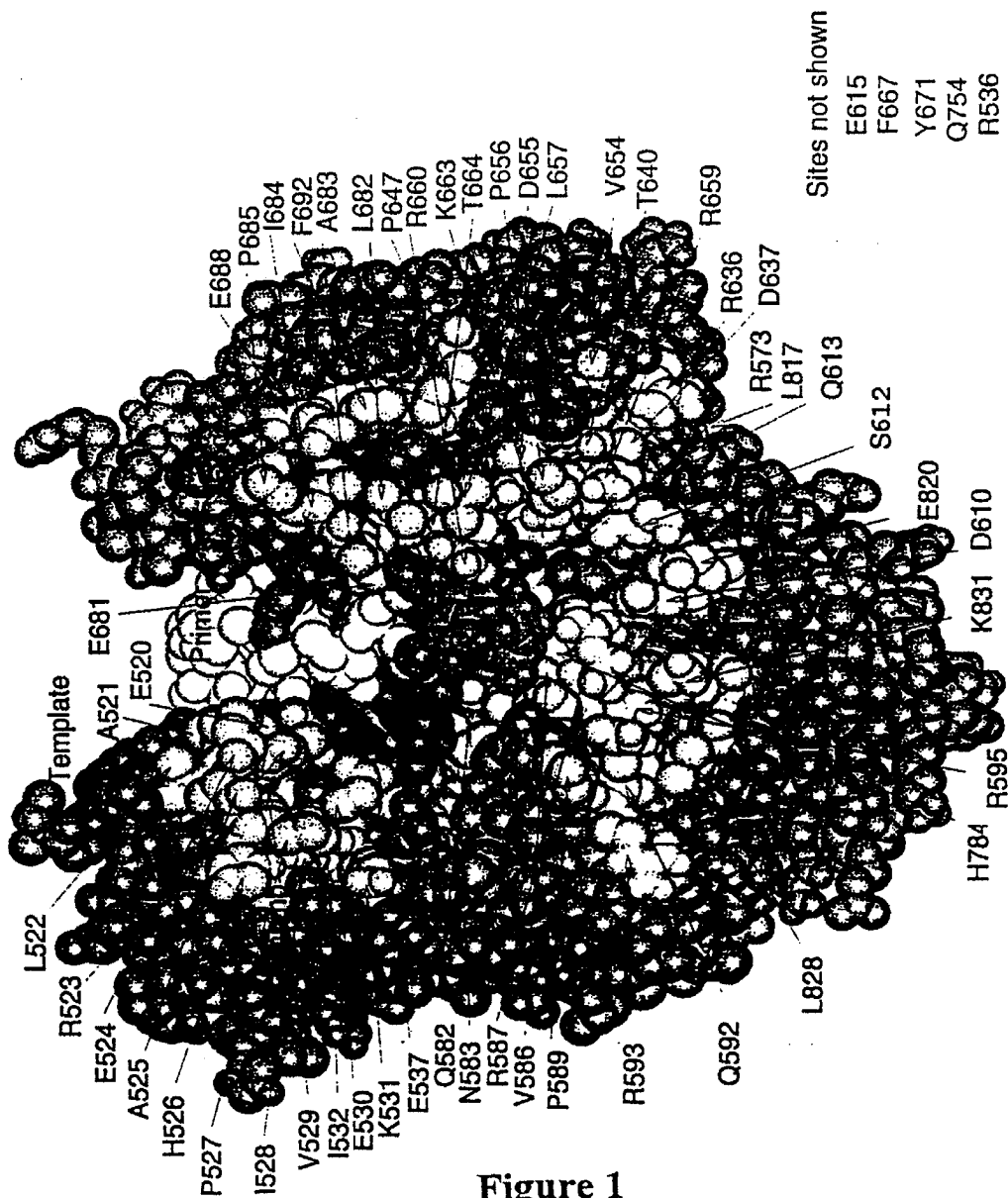
15

13. A method according to claim 12, wherein the primed template is a primed template in a chain termination sequencing reaction.

14. A method according to claim 12, wherein the primed template is a primed template in a polymerase chain reaction.

20

15. A kit for fluorescently labeling a polynucleotide, the kit comprising a DNA polymerase according to claim 1 and a fluorescently labeled nucleotide.



SUBSTITUTE SHEET (rule 26)

Next Nucleotide Effect Assay

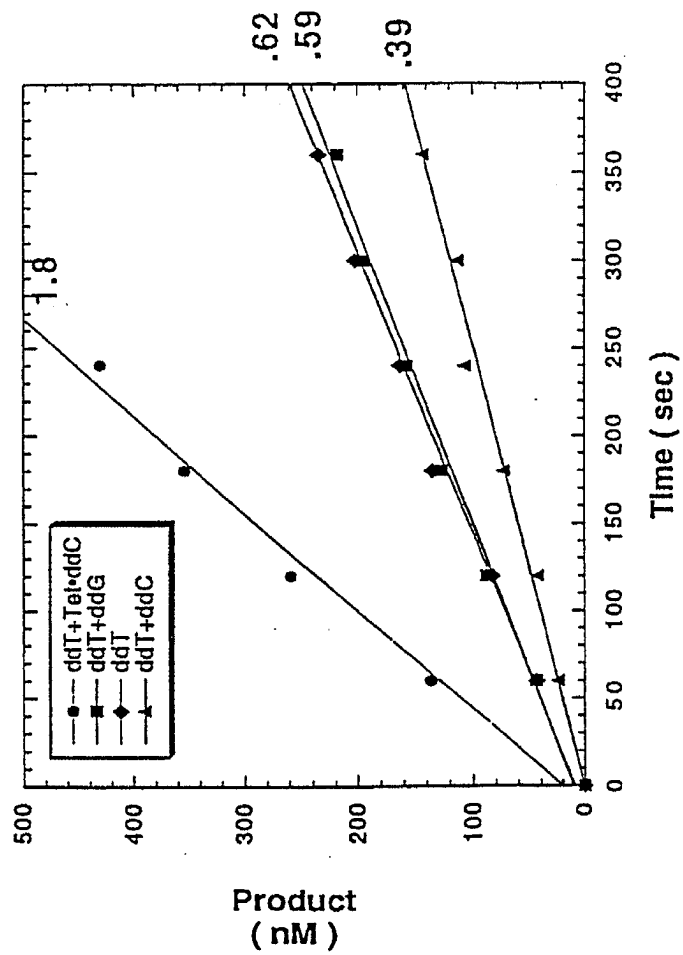
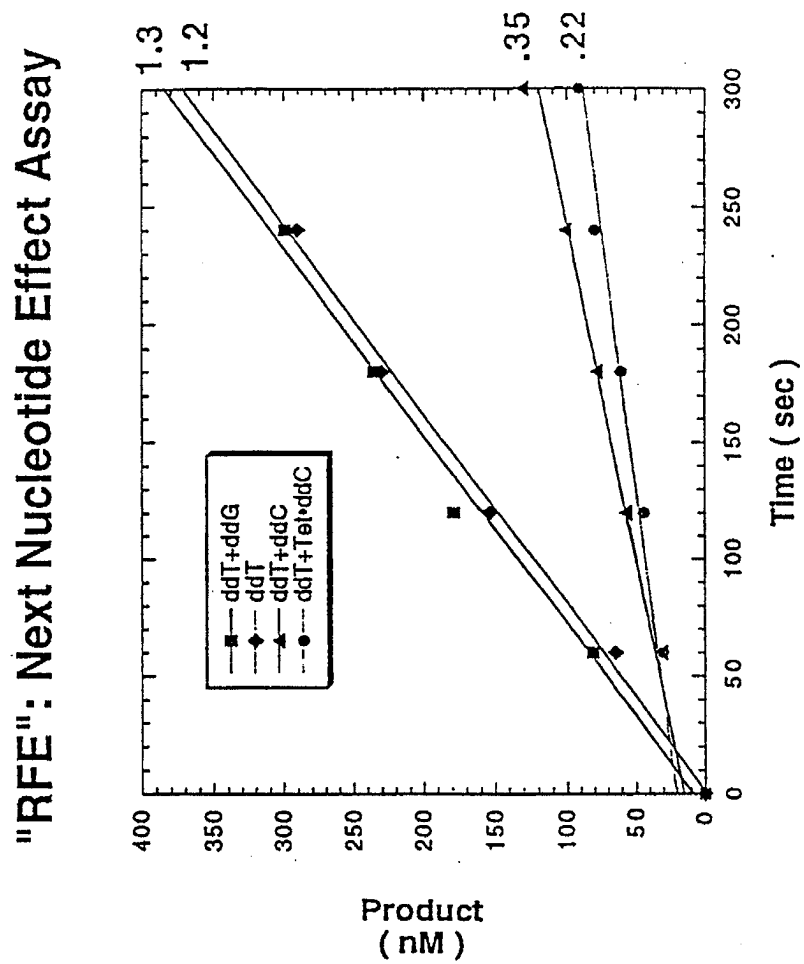


Figure 2

2/4

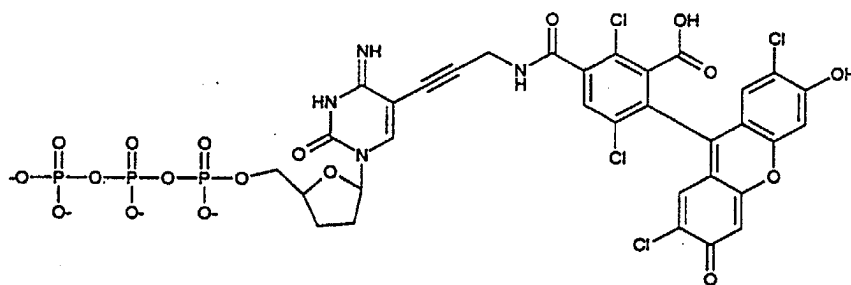
SUBSTITUTE SHEET (rule 26)

**Figure 3**

3/4

SUBSTITUTE SHEET (rule 26)

TET-2 ddCTP

**Figure 4**

4/4

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 98/05095

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/54 C12N9/12 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PARKER L. T. ET AL.: "AmpliTag DNA polymerase, FS dye-terminator sequencing: analysis of peak height patterns" BIOTECHNIQUES, vol. 21, no. 4, October 1996, pages 694-699, XP002067706 see abstract; tables 1,2 see page 695, paragraph 2 - page 696, paragraph 1 see page 699, paragraph 2 --- -/--	1,2,8-15

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

15 June 1998

Date of mailing of the international search report

25/06/1998

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Oderwald, H

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/05095

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SUZUKI M ET AL.: "Random mutagenesis of Thermus aquaticus DNA polymerase I: concordance of immutable sites in vivo with the crystal structure" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA, vol. 93, September 1996, pages 9670-9675, XP002067707 see abstract; figures 2-4 see page 9671, paragraph 4 - page 9672, paragraph 6 ----	1-5,8-11
X	REEVE M A ET AL: "A NOVEL THERMOSTABLE POLYMERASE FOR DNA SEQUENCING" NATURE, vol. 376, 31 August 1995, page 796/797 XP000606193 see the whole document ----	1,2,8-15
X	EP 0 727 496 A (HARVARD COLLEGE) 21 August 1996 see abstract; figures 3,6; examples 1,7,11,20; table 2 ----	1,2,8-15
X	KALMAN L. V. ET AL.: "Thermostable DNA polymerases with altered discrimination properties" GENOME SCI TECHNOL, vol. 1, 1995, page 42 XP002067708 see abstract A-14 see the whole document ----	1,2,8-15
P,X	EP 0 823 479 A (HOFFMANN LA ROCHE) 11 February 1998 see abstract; examples 1,6; table 1 see the claims see page 8, paragraph 3 see page 2, line 30 - page 4, line 4 ----	1-3,8-15
P,X	VOSS H ET AL: "AUTOMATED CYCLE SEQUENCING WITH TAQUENASE TM: PROTOCOLS FOR INTERNAL LABELING, DYE PRIMER AND "DOUBLEX" SIMULTANEOUS SEQUENCING" BIOTECHNIQUES, vol. 23, no. 2, August 1997, pages 312-318, XP000698417 see the whole document -----	1,2,8-15

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/05095

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0727496 A	21-08-1996	US 5614365 A	25-03-1997
		DE 29513622 U	19-10-1995
		DE 29513639 U	19-10-1995
		JP 2691149 B	17-12-1997
		JP 8205874 A	13-08-1996
		AT 143057 T	15-10-1996
		AU 4193396 A	06-05-1996
		CA 2201885 A	25-04-1996
		DE 69400567 D	24-10-1996
		DE 69400567 T	06-02-1997
		DE 655506 T	28-09-1995
		DK 655506 T	10-03-1997
		EP 0655506 A	31-05-1995
		ES 2072238 T	16-07-1995
		FI 971611 A	13-06-1997
		HU 77037 A	02-03-1998
		WO 9612042 A	25-04-1996
		ZA 9508761 A	13-05-1996
EP 0823479 A	11-02-1998	AU 3319797 A	12-02-1998
		JP 10066588 A	10-03-1998
		NO 973595 A	09-02-1998
		PL 321478 A	16-02-1998